Genetic Variation in Genes Associated with Milk Production Traits and Fatty Acids Composition in New Zealand Dairy Cattle

---

A thesis
submitted in partial fulfilment
of the requirements for the Degree of
Doctor of Philosophy

at
Lincoln University
by
Ishaku Lemu Haruna

Lincoln University
2020

Abstract of a thesis submitted in partial fulfilment of the
requirements for the Degree of Doctor of Philosophy.

Genetic Variation in Genes Associated with Milk Production Traits and Fatty Acids Composition in New Zealand Dairy Cattle

by
Ishaku Lemu Haruna

Abstract

The myostatin gene (*MSTN*), alternatively known as the growth and differentiation factor 8 gene (GDF8), encodes the myostatin protein (MSTN). It has pleiotropic effects, and it is expressed in skeletal muscle as well as the mammary gland. Its expression has been associated with increased skeletal muscle mass, and decreased adipogenesis as a result of the reduced secretion of leptin.

Leptin is a product of the obese (ob) gene. It is a 16-kDa protein hormone mainly secreted from white adipose tissue and it regulates feed intake, energy partitioning and metabolism, as well as lactogenesis.

Using Polymerase Chain Reaction (PCR) amplification, coupled with Single Strand Conformation Polymorphism (SSCP) analyses and subsequent nucleotide sequencing, an extended region of the bovine *MSTN* and leptin (*LEP*) genes were investigated for genetic variations in a variety of cattle breeds from New Zealand and Nigeria.

Five regions of *MSTN* were investigated. These included exons 1, 2 and 3, and parts of introns 1 and 2; and they were studied in 883 cattle of varied breeds from New Zealand (Hereford, Angus, Charolais, Simmental, Red Poll, South Devon, Shorthorn, Murray Grey, cross-bred Holstein-Friesian × Jersey cattle (NZ HF × J-cross, or Kiwicross^{TM}) and from Nigeria (Sokoto Gudali, Red Bororo, White Fulani and cross bred Holstein-Friesian × White Fulani). Eight PCR-SSCP banding patterns were observed in the intron 1 region amplified, four in the intron 2 region amplified and none in the exon regions. A total of seventeen single-nucleotide substitutions (nine in intron 1 and eight in a region spanning the intron 2 - exon 3 boundary) and one nucleotide deletion were detected.

For the leptin gene (*LEP*), three regions (intron 1/exon 2, part of intron 2 and the entire exon 3) were examined in 657 cattle from a variety of breeds farmed in New Zealand and Nigeria. These included NZ Hereford, Angus, Shorthorn, and NZ HF × J-cross cows; and the Nigerian Sokoto Gudali, Red Bororo, White Fulani, and cross-bred Holstein-Friesian × White Fulani cattle. Four PCR-SSCP patterns were detected in the intron 1/exon 2 region with four nucleotide sequence variations, three

patterns in intron 2 with three sequence variations, and three patterns in exon 3 with five nucleotide sequence variations.

The effect of intron 1 *MSTN* variation on milk production traits and the composition of milk FA were investigated. The saturated fatty acids (SFAs) were grouped into three groups based on the length of their carbon chain; Short chain fatty acid (SCFA: C4:0 to C8:0), medium chain fatty acid (MCFA: C10:0 to C14:0) and long chain fatty acid (LCFA: C15:0 to C24:0). The unsaturated fatty acids (UFAs) with only one double bond (monounsaturated fatty acid: MUFA) ranged from C10:1 to C22:1 and were grouped as MUFA, whereas those with two or more double bonds (polyunsaturated fatty acid: PUFA) ranged from C18:2 to C22:5 and were grouped as PUFA.

Using General Linear Mixed-effect Model (GLMM) analyses, single variant presence/absence models revealed the presence of variant *D* to be associated with reduced levels of C8:0, C10:0, C12:0, C12:1, and the grouped MCFA level; but an increased level of C16:1 *cis*-9. Variant *B* was associated with decreased C15:0 *iso* level, whereas the presence of *C* was associated with increased levels of C20:3 *cis*-8, 11, 14 and C22:1, *trans*-13. Only the NZ HF × J-cross cows with the following *MSTN* genotypes; *AA* (n = 151), *AB* (n = 92), *AC* (n = 53) and *AD* (n = 65) were studied in the milk FA genotype association analyses; with the remaining genotypes, *AE* (n = 15), *BB* (n = 7), *BC* (n = 7), *BD* (n = 17), *CC* (n = 6) *CD* (n = 10) and *DD* (n = 7) having frequencies less than 5% each, and not being analysed. These genotype model results were consistent with the single variant presence/absence models, with genotype *AD* being associated with reduced C10:0, C12:0, and C12:1 levels.

For the leptin gene, the effect of exon 3 variation on milk production traits and FA composition in NZ HF × J-cross dairy cows were investigated. Association studies using GLMMs revealed the presence of variant A_3 (the most common variant) decreased the levels of C15:1 C22:0, C24:0, C18:1 trans-11, C18:1 *trans*-9, *cis*-12 and C18:1 all *trans*. Variant B_3 was revealed to be associated with reduced C6:0, C11:0 level, and C20:0 level, but increased C17:0 *iso*, C24:0 and C10:1 index. Variant C_3 was associated with decreased C13:0 *anteiso*. The following genotypes; A_3A_3 (n = 73), A_3B_3 (n = 176) and A_3C_3 (n = 52) were examined for milk FA composition, whereas B_3B_3 (n = 11) and C_3C_3 (n = 3) were excluded from the analyses due to their low frequency of occurrence. In the genotype model, relative to the A_3A_3 genotype, the A_3B_3 genotype was associated with decreased levels of C8:0, C10:0, C11:0, C13:0 and grouped MCFA, but increased C24:0. Genotype A_3C_3 was associated with decrease grouped MCFA level and C10:0, C11:0 and C13:0 levels relative to the A_3A_3 genotype.

These associations in NZ HF × J-cross cows suggest that variation in bovine *MSTN* and *LEP* could be explored for increasing the concentration of UFAs and decreasing the concentration of SFAs in milk.

Keywords: Myostatin gene (*MSTN*), leptin gene (*LEP*), genetic variation, coding and non-coding sequences, PCR-SSCP, cattle, genotyping, nucleotide variation, gene marker, milk production, fatty acids.

Publications arising from this thesis

Papers

Haruna, I.L., Ekegbu, U.J., Ullah, F., Amirpour-Najafabadi, H., Zhou, H., and Hickford, G.H.J. (2020). Genetic variations and haplotypic diversity in the Myostatin gene of New Zealand cattle breeds. *Gene*, 740, https://doi.org/10.1016/j.gene.2020.144400.

Haruna, I.L., H., Zhou, H., and Hickford, G.H.J. (2020). Short commentary on: Absence of genetic variation in the coding sequence of Myostatin gene (*MSTN*) in New Zealand cattle breeds. *Journal of Data Mining in Genomics and Proteomics*. 11:222. DOI:10.35248/2153-0602.20.11.222

Haruna, I.L., Hadebe, S.A., Oladosu, O.J., Mahmoud, G., Zhou, H., and Hickford, G.H.J. (2020). Identification of novel nucleotide sequence variations in an extended region of the bovine leptin gene (*LEP*) across a variety of cattle breeds from New Zealand and Nigeria. *Archives Animal Breeding*, 63, 241-248. https://doi.org/10.5194/aab-63-241-2020

Haruna, I.L., Li, Y., Ekegbu, U.J., Amirpour-Najafabadi, H., Zhou, H., and Hickford, G.H.J. (2020). Associations between bovine *MSTN* variants and milk FAs in New Zealand Holstein-Friesian × Jersey-cross cows. *Animals*, 10, 1447; doi: 10.3390/ani10091447

Haruna, I.L., Li, Y., Zhou, H., and Hickford, G.H.J. (2020). Effects of bovine leptin gene variation on milk traits in New Zealand (NZ) Holstein-Friesian × Jersey (HF × J)-cross dairy cows (Kiwicross^{TM} cows). *New Zealand Journal of Agricultural Research;* doi.org/10.1080/00288233.2020.1838570

Haruna, I.L., Zhou, H., and Hickford, G.H.J. (2020). Variation in bovine leptin gene affects the composition of milk fatty acids in New Zealand (NZ) Holstein-Friesian × Jersey (HF × J)-cross dairy cows (Kiwicross^{TM} cows). *Archives Animal Breeding* (under review).

Conference presentations and posters

Haruna I.L., (2019). Associations between bovine MSTN variants and milk FA composition in New Zealand Holstein-Friesian × Jersey-cross cows. 37th International Society of Animal Genetics conference, 7-13 July Lleida, Spain.

Haruna I.L., (2018). Genetic variation and haplotypic diversity in the myostatin gene of New Zealand cattle breeds. Lincoln University Postgraduate conference, 20 – 24th September 2018 at Lincoln University, Canterbury, New Zealand.

Nucleotide sequences submitted to NCBI GenBank

Bovine MSTN intron 1 region

Myostatin gene (*MSTN*) intron 1, variant *A* sequence: MK353501

Myostatin gene (*MSTN*) intron 1, variant *B* sequence: MK353502

Myostatin gene (*MSTN*) intron 1, variant *C* sequence: MK353503

Myostatin gene (*MSTN*) intron 1, variant *D* sequence: MK353504

Myostatin gene (*MSTN*) intron 1, variant *E* sequence: MK353505

Myostatin gene (*MSTN*) intron 1, variant *F* sequence: MK353506

Myostatin gene (*MSTN*) intron 1, variant *G* sequence: MK353507

Myostatin gene (*MSTN*) intron 1, variant *H* sequence: MN510780

Bovine MSTN intron 2 region

Myostatin gene (*MSTN*) intron 2, variant A_4 sequence: MK353508

Myostatin gene (*MSTN*) intron 2, variant B_4 sequence: MK353509

Myostatin gene (*MSTN*) intron 2, variant C_4 sequence: MK353510

Myostatin gene (*MSTN*) intron 2, variant D_4 sequence: MN510781

Bovine leptin exon 2 region

Leptin gene (*LEP*) exon 2, variant A_1 sequence: MN082388

Leptin gene (*LEP*) exon 2, variant B_1 sequence: MN082389

Leptin gene (*LEP*) exon 2, variant C_1 sequence: MN082390

Leptin gene (*LEP*) exon 2, variant D_1 sequence: MN082391

Bovine leptin intron 2 region

Leptin gene (*LEP*) intron 2, variant A_2 sequence: MN069837

Leptin gene (*LEP*) intron 2, variant B_2 sequence: MN069838

Leptin gene (*LEP*) intron 2, variant C_2 sequence: MN069839

Bovine leptin exon 3 region

Leptin gene (*LEP*) exon 3, variant A_3 sequence: MN119553

Leptin gene (*LEP*) exon 3, variant B_3 sequence: MN119554

Leptin gene (*LEP*) exon 3, variant C_3 sequence: MN119555

Acknowledgements

I am grateful to the New Zealand Government, through the Ministry of Foreign Affairs and Trade, for considering me worthy of a New Zealand Development Scholarship, without which, this degree wouldn't have been possible.

My deepest appreciation goes to my supervisor; Professor Jon Hickford. Jon, as he would want me to call him, has been very supportive both on and off my PhD journey. His door is always open to listen and answer my never-ending questions, in addition to giving me the support I needed to establish and consolidate my networks-a recipe for success in my career.

I am also very thankful to my co-supervisor, Dr Huitong Zhou. Huitong's attention to detail has helped refined my experimental and analytical skills. I owe thanks to Mimi for her kindness and patience in introducing me to the basics of experimental protocols at the gene marker lab. I'll also like to thank Freeman and Andrea for ensuring availability of consumables in the lab, and for being kind in answering questions and giving suggestions during my experiments.

To my awesome colleagues at the gene marker lab, I am deeply grateful for taking this journey along with you: Jane, thank you for your sacrifice in putting me through those experimental and bioinformatics procedures, and for coming up with the "Jane's protocol". Thank you Yunhai and Hameed for topping up my buffer during experiments.

To my great friends; Christopher, Sibusiso, John, Samantha, Oveka, Lughano and members of the African and Caribbean club, thank you for providing me that balance between academics and social life.

I am indeed thankful to my family for the moral and spiritual support. To my beloved wife Washina, thank you for your understanding and the sacrifices you had to make while I pursue my dreams. To my beautiful kids; Ishaku Jnr, Semaiah and Jahleel, get ready for "hide n seek", daddy's coming home. To my mother (Anty), thank you for constantly praying for my success.

In the end, I'll like to dedicate this work to God almighty for his divine grace, wisdom and understanding, which has guided me throughout this journey.

Abbreviations

%	Percent
°C	Degrees Celsius
µg	Microgram
µL	Microlitre
µm	Micrometre
µM	Micromolar
A	Adenine
bp	Base pair
C	Cytosine
C-terminal	Carboxy terminal
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
dGTP	Deoxyguanosine triphosphate
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
dTTP	Deoxythymidine triphosphate
eBV	Estimated Breeding Value
EDTA	Ethylene diamine tetra acetic acid
g	Gram
G	Guanine
GLM	General linear model
GLMM	General Linear Mixed-effects Model
h	Hour
HGNC	Human Genome Nomenclature Committee
HGVS	Human Genome Variation Society
SSCP	Single strand conformation polymorphism
T	Thymine
HWE	Hardy-Weinberg Equilibrium
Kb	Kilobase
kDa	KiloDalton
kg	Kilogram
LEP	Leptin gene
mg	Milligram
micron	Micrometer
min	Minute
mL	Millilitre
mm	Millimetre
mM	Millimolar
MSTN	Myostatin
mRNA	Messenger RNA
ng	Nanometre
nm	nanometre
nt	nucleotide
NZ	New Zealand
Ob	Obese gene
PCR	Polymerase chain reaction
QTL	Quantitative trait loci
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
rpm	Revolutions per minute
SE	Standard error

SNP	Single nucleotide polymorphism
Taq	*Thermus aquaticus*
TBE	Tris-borate-EDTA buffer
TE	Tris-EDTA buffer
Tris	Tris (hydroxy methyl) aminomethane
UHS	Ultra high sulphur
U	Unit
UTR	Untranslated region
UV	Ultraviolet
V	Volt
Wt	Weight

Amino acids abbreviations

A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic acid
E	Glu	Glutamic acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

Table of Contents

List of Tables

List of Figures

Chapter 1

Literature Review

The word 'dairy' can be considered synonymous with New Zealand's economy. In 2010 the dairy industry contributed about 2.8% of GDP (approx. NZ$5.0 billion) to the economy, this comprising 26% of New Zealand's total export goods (Chris, 2010), while by 2016 data from the New Zealand Institute of Economic Research (NZIER) revealed an increased GDP contribution of 3.5% (approximately NZ$7.8 billion) (John, 2017). This trend continued, as by 2017, the NZIER's report on the dairy sector's contribution rose to NZ$8.2 billion (Destremau and Siddharth, 2018), and on a global scale, New Zealand produced 3% of all milk in the world in the 2016/2017 season (DairyNZ Quickstats, 2017). More recently, it was revealed that dairy farming contributed 31% of the total goods exported from New Zealand in the year 2018 (DairyNZ Quickstats, 2019).

With these data, and taking into account the fact that milk of dairy cows is of high nutritional value to humans, and therefore sought after globally; it is important to examine the environmental and genetic factors that could affect milk production for the purpose of improving the quality and quantity of the milk produced.

This literature review consists of (i) an overview of factors affecting gross milk production traits in New Zealand dairy cows (ii) a description of the genetic challenges confronting the dairy industry in New Zealand (iii) the transition to genomic selection for improved milk production and FA composition with a focus on a candidate gene approach and (iv) a discussion of some selected genes (myostatin (*MSTN*) and leptin (*LEP*)) that potentially affect milk production and the composition of milk FAs.

The location and molecular organisation of *MSTN* will be highlighted, and its expression in the mammary glands of mouse and pigs will be discussed in relation to its effect on milk production. Likewise, the structure of the leptin gene will be discussed, its regulatory and reproductive functions will be highlighted, and its role in nutritional status will also be addressed. Finally, genetic variation in both the *MSTN* and *LEP* genes and its association with milk traits of economic interest will be reviewed.

1.1 Factors affecting milk traits and milk fat composition in New Zealand dairy cows

There are several factors that can influence gross milk traits such as milk yield, overall protein and fat content, and the constituent make-up of the milk fat, milk protein and other milk components. Some of these factors include; variation in climate, types of diet, feeding system, breed selection, cow age, and

the different stages of lactation. It is important to note that, each of these factors could control more than one milk-related trait, hence investigating the effect(s) of these factors will by definition require a multivariate analysis.

From previous reports, factors such as breed differences, lactation stage, diet, animal body-condition and the environment have been found to affect the production of milk fat (Stelwagen, 2011), whereas the effects of dairy production system, feeding regime, herd, cow parity and stage of lactation have been suggested to affect milk FA composition (Mele *et al.*, 2016).

1.1.1 The effect of pasture-based diets on composition of milk FA

In New Zealand, the dairy production system is in large part pasture-based, and the most favoured pasture consists of perennial ryegrass (*Lolium perenne*) and white clover (*Trifolium repens*) (James *et al.,* 1999). The popularity of this pasture is based on it being a low cost, high quality nutrient source (Penno *et al.,* 2007) with minimal need for additional feed supplementation. In the 2016-2017 season, the report of the DairyNZ Economic Survey (DairyNZ Quickstats, 2018a) revealed that 82% of all feed eaten by dairy cows in New Zealand was pasture, but with this including the preserved feeds of hay, baleage and silage.

One of the advantages of such a production system is that it increases the amount of PUFA and conjugated linoleic acids (CLA) in the milk as suggested by Chilliard *et al.* (2001) and Dewhurst *et al.* (2006). These fats are considered to be of benefit to human health (Chilliard *et al.,* 2001).

1.1.2 The effect of lactation stage on milk fat production and composition

One of the unique features of the New Zealand pastured-based dairy production system is its seasonality, and accordingly the milked cows are usually at a very similar stage of lactation. This is notable because milk fat production and composition are often affected by the different stages of lactation, and mainly as a consequence of the change in the cow's body-condition as lactation progresses, particularly in terms of energy source and utilisation. The findings of Strucken *et al.* (2015) suggest that cows in the early stage of lactation (first 60 days in milk (DIM) – the first trimester) are typically in a negative energy balance. This means, the energy the cow derives from feed intake cannot match the demand of high milk production. Consequently, an alternative source of energy is needed in order to offset this negative energy balance, and hence the cow's body fat reserves are mobilised to balance the deficit between feed intake, and energy expenditure on maintenance and milk production (Bauman *et al.,* 1980). However, as a result of using the cow's body fat, other biological pathways are affected, which results in a change in milk composition. Figure 1.1 illustrates the energy balance associated with the demands of high milk production in early lactation.

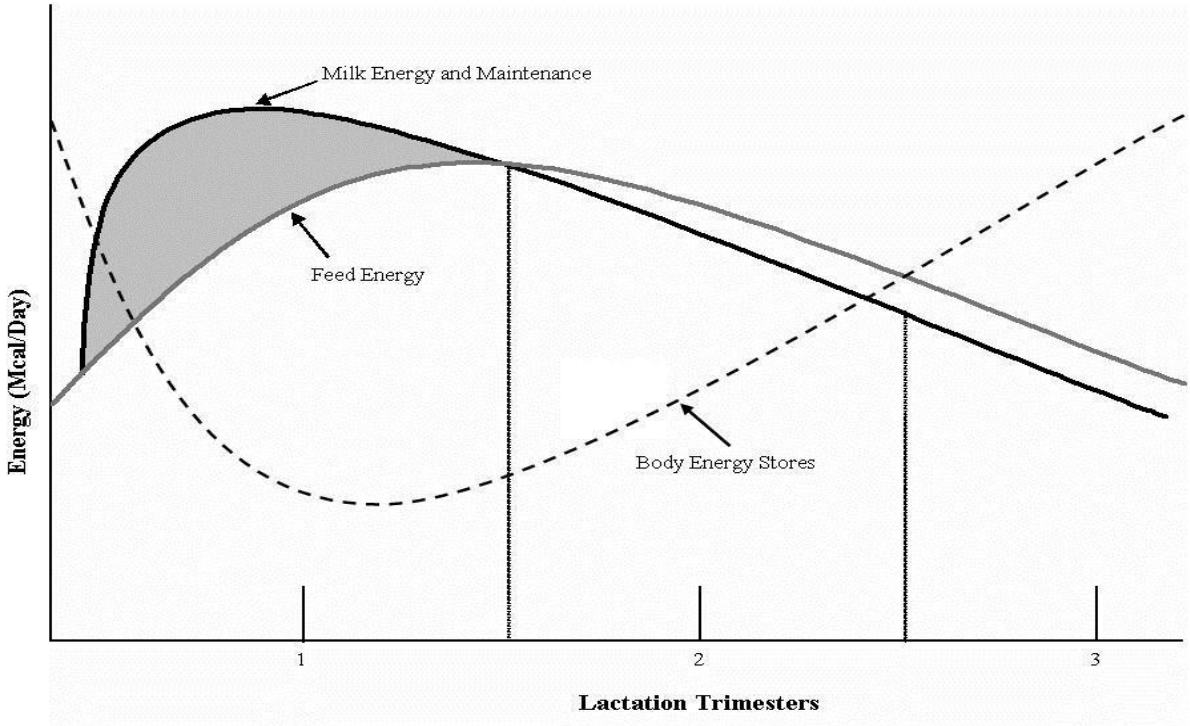


Figure 1.1 Archetypal model of the energy supply and requirements for milk production and maintenance during the lactation cycle of cows.

The black line represents the energy demand for milk production and maintenance, the grey line represents energy supply from dietary intake, the broken line represents energy from body fat reserves and the period of negative energy balance is indicated by the grey shaded area in approximately the first 60 days of lactation in the first trimester (https://www.influxlipids.com/single-post/2017/08/12/Feeding-High-Yield-Milk-Cows-While-Maintaining-Milk-Yield-And-Fertility).

In the mid- and late-lactation stages, the source of energy supply for milk fat production becomes feed intake and this leads to the *de-novo* synthesis of FAs. These stages are typically characterised by an increase in milk fat percentage. The composition of the milk fat also changes, because of the change in the ratio of *de-novo* or synthesised FA, to 'imported' fat.

Generally, cows in the first trimester of lactation (characterised by the negative energy balance period), tend to produce a higher level of SFA (mainly C16:0 and C18:0 FA), but when the period of negative energy balance ends, *de novo* synthesis of FA becomes the primary source of milk fat and typically the saturation ratio of milk FA changes after the 60th DIM . Also, in the mid- and late-lactation stages, most of the milk FA (except the C5, C15, branched and CLA *trans*-10, *cis*-12 FA) levels can undergo change (Stoop *et al.*, 2009).

In addition to the changes in the energy supply and requirements for milk production and maintenance during the lactation cycle (as illustrated in Figure 1.1), it is important to stress that inadequate feeding of cows can lead to shorter, lower yielding lactations, and subsequently increase calving interval. In this context, there appears to be an interrelationship between the stages of the lactation, cow live weight, dry matter intake (DMI) and overall milk yield. Other changes can also affect milk FA

content, including whether the cow is pregnant. Figure 1.2 contains a generalised model of the interrelationships between feed intake, milk yield and live weight for a Friesian cow with a 14-month inter-calving interval (360 d lactation), albeit in the New Zealand dairy system cows are typically mated again within weeks of the start of lactation and lactation has ceased at day 300 or before of lactation.

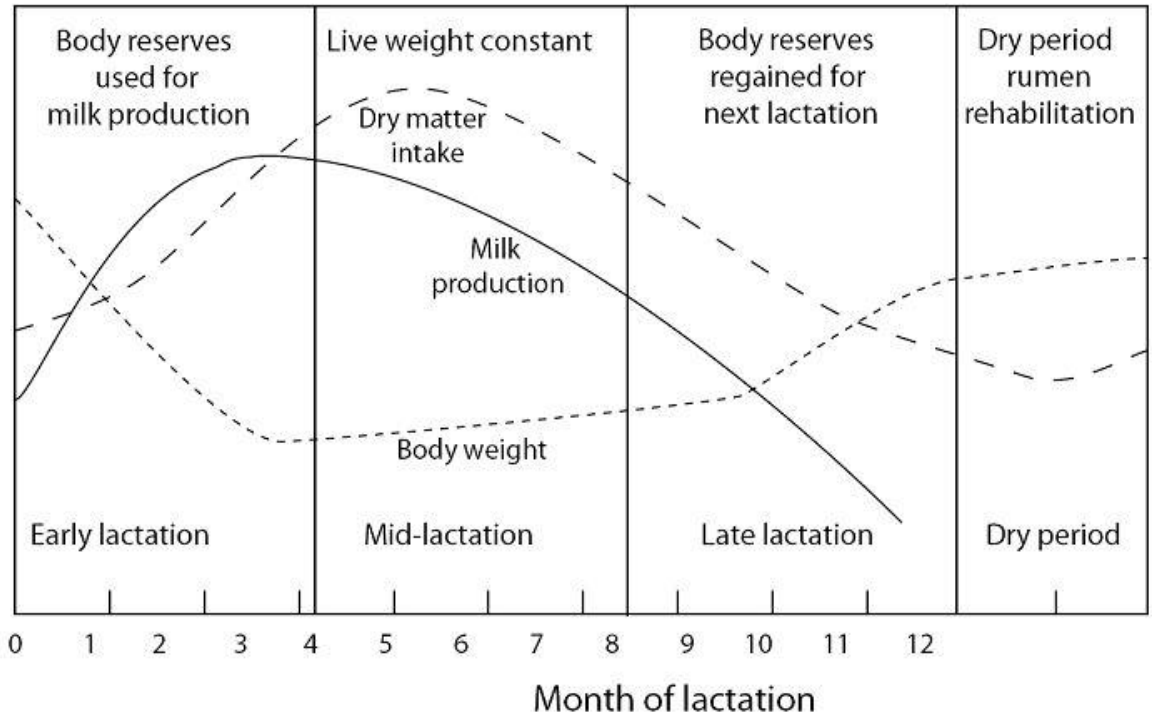


Figure 1.2 Generalised model of the changes observed in dry matter intake, milk yield and live weight in relation to the various stages of a cow's lactation.
The cow's body reserves are used up in the early stages of lactation, and by mid lactation the dry matter intake peaks, which enable the body reserve to be regained in the later stages of lactation (Adapted from Moran, 2015).

Lactation from calving to peak lactation

At the peak of lactation, milk yield is a good indicator of milk production for the year, with an extra kg per day at the peak being associated with up to an extra 200 kg/cow over the entire lactation period (Moran, 2015). However, there are a number of limiting factors affecting the proper feeding of cows in early lactation, with these affecting the ability to maximise the peak. One such factor is voluntary feed intake, which is an indicator of cow appetite. The appetite at calving is usually about 50 to 70 per cent of the maximum appetite at peak intake (Moran, 2015). This is usually because, during the dry period, the growing calf occupies space, thereby reducing the volume of the rumen, as well as the density and size of rumen papillae. After calving, it takes a while for the rumen to expand to capacity and the papillae to regrow. Accordingly, it is not until weeks 10-12 of lactation that appetite reaches its full potential (Moran, 2015).

Peak lactation and peak intake

This period is when cows tend to maintain weight. This is because, following peak lactation, the appetites of cows gradually increase until they are able to consume all the nutrients required for production (Moran, 2015). This is provided their diet is of a high-enough quality.

Mid and late lactation

This period is often associated with less demand for energy because of the decline in milk production. None-the-less, energy is still very important and often required because of pregnancy and the need to develop body-condition as an energy reserve for the next lactation. Generally, it is considered more efficient to try and improve the condition of the herd in late lactation rather than in the dry period (Moran, 2015).

The dry period

This is an important stage as maintaining (or increasing) body-condition during this period is critical to ensuring that cows have adequate body fat to fuel milk production and maintenance in the next lactation.

If calving occurs at a time when cows have adequate body reserves, they can cycle very soon after calving (Moran, 2015). However, if calving occurs when the cow is in poor condition, the production of milk in the early stage of lactation is affected due to the cow having inadequate body reserves (Moran, 2015). At the start of lactation, dietary energy can in fact be partitioned towards weight gain rather being made available from the desired weight loss (Moran, 2015). Consequently, high feeding levels in early lactation cannot 'make-up' for the effect of cows having poor body-condition at calving.

1.1.3 The effect of breed selection for increased milk solid (MS) production

The Breeding Worth (BW) index (www.dairynz.co.nz/animal/animal-evaluation/), is widely used as a selection tool for breeding in New Zealand dairy industry. It is based on an animal's genetic merit (constructed on the calculation of eBVs) and takes into account traits such as milk yield (measured in litres), milk fat production (measured in kilograms), milk protein production (measured in kilograms), cow live-weight, somatic cell score, fertility, body-condition score and residual survival. These are all traits that are considered valuable to the New Zealand dairy industry.

There are predominantly three types of dairy cattle in New Zealand; Jersey (J), Holstein-Friesian (HF) and NZ HF×J-cross (or Kiwicross^{TM}) cows. These breeds differ in terms of milk-related traits, especially in the composition of milk FAs. For example, milk from Jersey cows contains higher concentrations of some short- and medium-chain SFA, but lower concentrations of some UFA (Arnould and Soyeurt, 2009). Therefore, in the context of obtaining the preferred FA profile for human health benefit, these differences could be capitalized upon through the process of cross-breeding. In a study on the genetic gain of cross-breeding Holstein-Friesian and Jersey cattle breeds,

Lopez-Villalobos *et al.* (1996) showed that the cross-bred NZ HF × J cows in New Zealand have superior economic performance compared to either of the two parent breeds. Following this discovery, a genetic evaluation of the cross-bred cattle was initiated (Harris *et al.*, 1996) and an increase in the number of cross-bred HF × J cows was recorded as the proportion of NZ HF × J-cross cattle increased to 47.8% in the 2017/18 season (Figure 1.3).

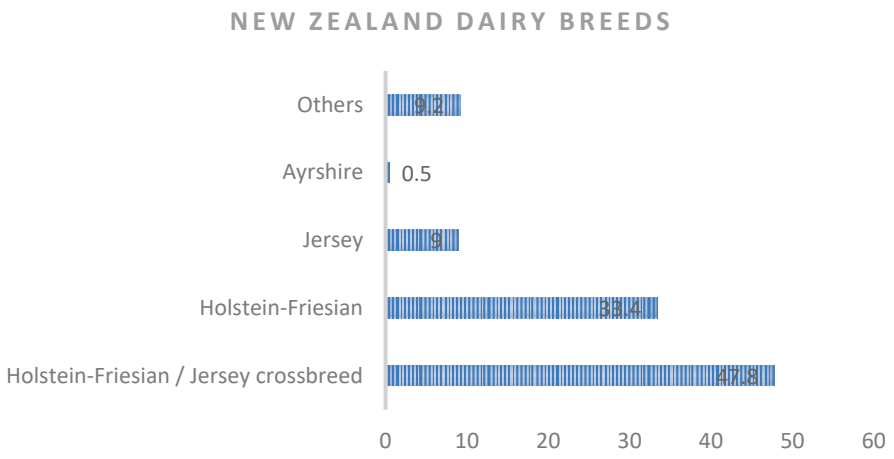


Figure 1.3 Breed/cross breakdown for dairy cows in New Zealand in the 2017/18 season (Adapted from DairyNZ, 2018b). The dominant breed in New Zealand is now the NZ HF x J-cross (Kiwicross™).

1.2 Genetic challenges confronting the dairy industry

One of the major genetic challenges confronting the dairy industry is how to improve the efficiency of the dairy production systems in terms of increasing milk yield and milk fat composition. In trying to ensure an efficient dairy production system, it is important to appreciate the fact that every animal has a fixed maintenance cost, and that this is a function of its size. This maintenance cost reflects the energy and nutrients that are required in order to maintain homeostasis. A larger cow requires a larger amount of feed to maintain its system, whereas a smaller cow requires relatively less feed. Therefore, if an animal is slow growing, then the amount of energy going into maintaining its system will be a greater proportion of the total energy requirement, and not necessarily be associated with an increased output in terms of milk yield. That is, faster growing animals will have a proportionally lower fixed maintenance cost and thus, improved efficiency in producing increased milk yield within a defined period of time.

1.2.1 Why genetic selection?

The need for genetic selection arises because of the disadvantages associated with the traditional animal breeding approaches that rely on selecting animals based on visual or phenotypic assessment to estimate their breeding value (BV) for any given trait. The traditional process can lack precision, be inefficient and time-consuming, with traits like milk production being hard to assess and evident only

after the animal reaches maturity. This results in a delay in verifying the success of the breeding approach.

Genetic selection approaches can increase selection accuracy. The approach entails the use of molecular genetic tools to understand the genes that underpin phenotypes of interest and more importantly, how this genetic merit is passed on from parents to offspring. It provides improvement over traditional phenotypic breeding systems, as it is focused on selecting animals based on their genetic potential for traits of interest and not their phenotypes.

1.2.2 The transition to genomic selection

The development of genomic selection began with the traditional method of selecting animals for breeding purposes. The traditional method of selecting animals, based on visual assessment, has been in existence since antiquity. This method entails the selection of animals that are more likely to pass their genetic traits of interest to their offspring based on their physical appearance or simple measurements of performance. Unfortunately, since the phenotypic appearance of an animal is both a function of its genetics and environmental influence, this approach has lots of pitfalls as the environmental factors that may be influencing an animal's phenotype, are not usually passed down to the offspring. This necessitates creation of a selection tool that takes into account or negates the effects of environmental factors.

The use of estimated Breeding Values (eBVs) is an improvement over the visual assessment approach, because of its ability to take into account how much of each animal's performance is due to its breeding merit and how much is due to the environment in which it has been raised. An eBV is a value which expresses the difference (+ or -) between an individual animal and the herd or breed benchmark to which the animal is being compared for any given trait. The eBV provides a measure of the breeding potential of an animal as a parent for a specific trait and takes into account performance data collected on that animal, known relatives, and its progeny, the relationships between performance traits (correlations), and the degree to which the trait is passed on from one generation to the next (its heritability).

In the dairy industry, the selection of sires is mainly based on the use of eBVs, but this method of selection has many disadvantages, which include: being time-consuming (the need to wait on daughters of potential sires to produce and for lactation information to be obtained). Also, this method is hampered by the relatively low heritability of fertility traits. The implication of this is that, a large number of daughters is required to be sure of the accuracy of selection.

In recent years, there has been a gradual shift from livestock selection based on phenotypic approaches and use of eBVs to 'genomic selection' approaches. In selecting for milk production for example, genetic selection is viewed as a long-term solution for genetic improvement in livestock species, as it entails the incorporation of more accurate approaches to increasing both the quantity and quality of

milk production. Meuwissen and Goddard (1996) reported that selection can be improved by integrating genetic information with information from phenotypic measurements. This will enable breeders to more cost effectively and accurately determine the genetic merit of an animal, through the creation of a genomic estimated breeding value (geBVs), and even in young animals (Muir, 2007).

Genomic selection will include the identification of sequence variation in key genomic regions, and analysis of associations between this genetic variation and phenotypic traits of interest. The identification of the genomic regions is undertaken either by using gene maps and a Quantitative Trait Loci (QTL) mapping approach, or by the study and use of candidate genes (Ovilo *et al.*, 2002), or by using both approaches together.

The quantitative trait loci approach (QTL)

The QTL technique is based on the use of DNA markers in genome or chromosome scans. In cattle, there have been several reports of genome and chromosome scans (Georges *et al.*, 1995; Heyen *et al.*, 1999; Lindersson *et al.*, 1998; Schrooten *et al.*, 2000). However, most of these reports only identified markers for specific traits of interest, but the gene(s) responsible for the effect remained unknown. One of the disadvantages of this technique is that it is difficult to figure out the causal gene for a specific marker, and this is because there are usually a large number of genes located in the vicinity of a marker. However, in a few QTL studies that have also used positional cloning techniques, the causal gene could be identified (Blott *et al.*, 2003; Grisart *et al.*, 2002).

The candidate gene approach

The name ‘candidate gene’ refers to a gene whose biological function might be controlling the occurrence of a certain phenotypic characteristic. The candidate gene approach entails the detection of variations in the nucleotide sequence of candidate gene(s) and an association study between these variations and economic traits of interest. Unlike the QTL approach, in the candidate gene technique, the function of the gene is known and potentially well understood, and the effects of variation in the gene is examined against traits of interest. Investigations involving the candidate gene(s) approach are carried out using the classical or reverse genetics techniques.

The forward genetics (classical genetics) approach is usually dependent on three key principles: identification of a particular phenotype of interest, then trying to understand why the phenotype exists, and finally, identifying the key gene controlling the phenotype. The limitation in this approach is that, it can result in the investigator(s) having a narrow view of the gene(s) in question, as it may not take into account the pleiotropic effects of the gene(s). Additionally, this approach is often focused on one or a small number of sequence variants or SNPs, which can lead to researchers having tunnel-vision. Additionally, taking into account the observation that there may be more than one gene influencing a trait, investigation of multiple genes using this approach might be time consuming (Andersson, 2001). Furthermore, caution must be taken in interpreting results obtained from this approach, as linkage

disequilibrium to the actual causative gene might affect the data (Ovilo *et al.*, 2002), albeit this problem also exists with QTL mapping approaches.

The reverse genetics approach ensures a more detailed understanding of the pleotropic effects of gene(s) associated with various traits. This approach is based on investigating the expression of the coding sequences of a gene in different tissues and examining how the protein product from these tissues can affect different traits of importance. This is a more elaborate way of understanding genes and their effects.

Overall, the selection of genes that are expressed in the mammary gland or underpin lactogenesis and adiposity is a sensible starting point for investigation of how genes may affect milk production traits and the composition of milk fat.

1.3 Genes affecting milk production and the composition of milk fat

While there are several metabolic pathways such as FA transport, FA synthesis, FA desaturation and release that can suggest the genes (or gene activities) that might contribute to variation in milk fat production and composition, the specific expression of a gene in the mammary gland and the effect of a gene or genes on feed intake are two other important clues that may indicate genes that affect milk production and milk fat composition. In this context, the myostatin gene *(MSTN*) and the leptin gene (*LEP*) were selected for investigation, the former because it was known to be expressed in the mammary gland and because of its role in muscle growth and carcass traits, and the latter because of its well recorded effect on appetite.

1.3.1 The myostatin gene (MSTN)

The gene *MSTN* has been located and characterised in a large number of species. In mice, it is located on chromosome 1 (McPherron *et al.*, 1997), and in sheep it is located on chromosome 2 (Clop *et al.*, 2006); whereas in pigs, horses and dogs it is located on chromosome 15, 18 and 37 as reported by Sonstegard *et al.* (1998), Lowe *et al.* (1997) and (NCBI GenBank Accession Number NW876304) respectively.

In cattle, the myostatin gene is located 3.1 cM (centiMorgan) from the centromeric region on chromosome 2 (BTA2), next to the microsatellite marker TGLA44 (Charlier *et al.*, 1995; Grobet *et al.*, 1997; Smith *et al.*, 1997) and in humans, *MSTN* is in a similar region to the bovine gene (Charlier *et al.,* 1995; Dunner *et al.,* 1997; Smith *et al.,* 1997).

The molecular structure of bovine *MSTN*

In cattle, molecular analysis has revealed that *MSTN* consists of three exons and two introns. Exons 1, 2 and 3 consists of 373, 374, and 381 nucleotides respectively, while introns 1 and 2 consist of 1,840

and 2,033 nucleotides respectively (Jeanplong *et al.*, 2001). The promoter region of bovine *MSTN* contains the androgen response elements (ARE) and myocyte enhancer factor 2 (MEF2) elements, in addition to the TATA and CAAT boxes (Du *et al.*, 2005). The bovine promoter region also has a high homology (95.8%) with ovine *MSTN* promoter, 86.9% with porcine *MSTN*, 80.2% with human *MSTN* and 67.7% with the mouse *MSTN* promoter regions (Du *et al.*, 2005).

Nucleotide sequence variation in the coding and non-coding regions of *MSTN*

It is now accepted that *MSTN* is highly sequence variable, with single nucleotide variations occurring approximately every 100bp through both the coding and non-coding sequences (McPherron and Lee, 1997; Grobet *et al.*, 1998; Dunner *et al.*, 2003; Bellinge *et al.*, 2005). The previous reports of naturally occurring nucleotide sequence variation in *MSTN* have typically focused on their effect on muscle related traits, and variation in the coding sequence has been extensively studied in various mammals, and linked with muscularity in mice (McPherron and Lee, 1997), humans (Schuelke *et al.,* 2004), sheep (Clop *et al.,* 2006; Han *et al*., 2015; Hickford *et al.*, 2010), dogs (Mosher *et al*., 2007), and cattle (Dunner *et al*., 2013; Grisolia *et al*., 2009; Grobet *et al.,* 1998; McPherron and Lee, 1997). The variation occurring in the coding regions of bovine *MSTN* have been revealed to affect the amino acid sequence and underpin some of the so-called 'double-muscled' phenotypes. This variation includes: c.821_831del11, c.419_421del7ins10, p.Q204X, p.E226X, p.E291X, p.C313Y, p.F94L, p.S105C, and p.D182N (Dunner *et al.*, 2003; Esmailizadeh *et al.*, 2008; Grisolia *et al.*, 2009; Grobet *et al.*, 1998; Grobet *et al.*, 1997; Kambadur *et al.*, 1997).

Unlike the coding sequences, the non-coding regions of the *MSTN* have not been accorded that much scientific scrutiny. A possible explanation is because the non-coding sequences were initially thought to neither play a role in, nor affect gene expression. This archaic and naive view is now changing, following the findings of Chung *et al.* (1989) and Buchman *et al.* (1988) among others, who suggest that introns can have a large effect on gene expression in some species. This view was also shared by Bellinge *et al.* (2005), whose findings highlighted that variations occurring in the non-coding sequences of *MSTN* could have an effect on gene expression because the disruptive variations identified in the coding regions do not account for all of the inherited phenotypic variations in double-muscled cattle. In an *in-vitro* experiment involving the green fluorescent protein (GFP), He *et al.* (2010) transformed C2C12 cell-lines with a transgene construct that contained part of the bovine *MSTN* promoter (pMD-MSTNPro) and another construct containing the first intron of bovine *MSTN* (pMD-Intron1), along with a reporter gene (Green Fluorescent Protein; GFP). They observed an increase in fluorescence and the number of fluorescence positive cells and concluded that the presence of intron 1 of bovine *MSTN* improved the expression of GFP in the transformed cells.

Taken together, it could be concluded that any investigation of *MSTN* variation and its effect on milk traits, should include analysis of nucleotide sequence variation in both coding and non-coding sequences.

Myostatin expression

The transcript expressed from the coding region of the MSTN gene codes for a protein of 375 amino acids in length (Jeanplong *et al.*, 2001). The gene has pleiotropic effects, and its expression has been associated with increased skeletal muscle mass and decreased adipogenesis, as a result of the reduced secretion of leptin (McPherron *et al.*, 1997; Whittemore *et al.*, 2003; Mendias *et al.*, 2008).

Previous reports have established that MSTN is expressed both in skeletal muscle ((McPherron *et al.*, 1997) and the mammary gland (Manickam *et al.*, 2008). In the latter context, this thesis will focus on the expression of MSTN in the mammary gland, but the vast majority of the literature about MSTN has been narrowly focused on its role in meat production. This is not to say that this is unimportant, but this narrow focus tends to downplay the question around what else MSTN may be influencing in other tissues. The forward genetics argument has dominated the narrative, and at the expense of any thinking that *MSTN* might be a pleiotropic gene.

The mammary gland

The complexity of the function of the mammary gland is unquestionable, as both structural and functional changes occur throughout the various stages of its development (Richert *et al.*, 2000). Noticeable physical changes such as elongation and branching of the milk ducts mainly occurs after the onset of puberty, while the proliferation of mammary epithelial cells take place during pregnancy. The mammary gland only becomes fully differentiated and functional in terms of milk production only during lactation (Robinson *et al.*, 1995; Robinson *et al.*, 1999).

These phases of mammary gland development are often regulated by several hormones and growth factors (Hennighausen and Robinson, 2001). In mice, members of the transforming growth factor-ß (TGF-ß) superfamily of signal transduction proteins have been found to play a key role in mammary gland development (Robinson *et al.*, 1991). TGF-ßs are secreted signalling proteins whose main function is to regulate cellular proliferation and differentiation. To date, over 30 members of the TGF-ß family have been described, and they are similar in structure; but each member has functionally diverse effects on tissues (Massague, 1998). Myostatin is a member of the TGF-ß family (McPherron *et al.*, 1997).

The Expression of *MSTN* in the mammary gland of mice

Temporal and spatial expression of MSTN has been found in the mouse mammary gland during both pregnancy and lactation (Manickam *et al.*, 2008). In their report, Manickam *et al.* (2008) demonstrated that its expression was predominantly associated with epithelial cells and that its expression was inversely correlated with the differentiated state of the gland during pregnancy. Their study also highlighted distinct temporal profiles of both myostatin mRNA and protein levels in the mammary gland. The highest level of mRNA observed corresponded with the periods of maximal ductal growth, and that this decreased as pregnancy progressed and decreased to minimal levels at the beginning of lactation.

Although MSTN shows a similar profile of expression to the other members of the TGF-ßs during the mammary gland development of mice (Robinson *et al.*, 1991), there seem to be redundancy in the TGF-ß proteins active in the mammary gland. A possible explanation could be because expression of multiple TGF-ßs ensures that the mammary gland fully differentiates to maximise milk production. In this context, the expression of MSTN in mammary gland has only been detected in multiparous animals; pig (Ji *et al.*, 1998) and mouse (Manickam *et al.*, 2008) so far, where presumably milk production must be optimal. Another possibility could be, expression of multiple TGF-ßs might be dependent on the role of a common transcriptional activator with the expense of superfluous TGF-ß activity tolerated by the animal as suggested by (Manickam *et al.*, 2008). This is conceivable, considering MSTN is synthesised in a precursor form as described by McPherron *et al.* (1997).

In skeletal muscle cells, MSTN has been found to signal through both Smad2 and Smad3 (Zhu *et al.*, 2004). This suggests that MSTN may also signal through Smad2 and Smad3 in the mammary glands of mice, especially given that the presence and phosphorylation of Smad2 and Smad3 correlates with the processing of MSTN in the mammary gland (Manickam *et al.*, 2008). Also, since some transcription factors belonging to the Smad family have already been implicated in the development of mammary gland (Yang *et al.*, 2002), it is therefore logical to assume that expression of multiple TGF-ßs is merely acting to amplify Smad signalling. Similarly, just like the mammary gland, the expression of multiple TGF-ßs is also observed in the bone and has been implicated with various growth and differentiation phases during adult life (Cho *et al.*, 2002). Hence, it is possible that the expression of multiple TGF-ßs may be a common event during differentiation.

In skeletal muscle, MSTN has been established as a dominant-negative regulator of cell proliferation and differentiation. It suppresses expression of markers of muscle differentiation and inhibits muscle cell proliferation by inducing p21 expression (Thomas *et al.*, 2000). However, in the epithelial cells of mammary gland, the expression of MSTN neither inhibits the proliferation, nor induces an increase in p21 levels. There is accordingly a big difference between the downstream consequences of MSTN expression in mammary epithelial cells and skeletal muscle cells.

<u>The Expression of MSTN in mammary glands of pigs</u>

In pigs, MSTN is expressed in the tubule-alveolar secretory lobules of the lactating mammary gland (Ji *et al.*, 1998). This suggests that it may play a regulatory role in growth, development and/or metabolism in the gestational or lactating mammary gland. Taking into account the fact that milk from several species contains growth factors, including TGF-α and -ß2 **(**Ellis *et al.*, 1997)**,** this could suggest that the mammary gland secrets MSTN into the milk, where it can then play a regulatory role in the neonatal pig.

1.3.2 The Leptin gene (LEP)

The bovine leptin gene (*LEP*) previously known as *OB*, *OBS* and *LEPD* has been mapped to chromosome 4 (Pomp *et al.*, 1997) and it consists of three exons separated by two introns. Exon 1 and part of exon 2 (four nucleotides) are not translated, and only the remaining part of exons 2 and 3 are translated into the functional 16-kDa leptin protein of 146 amino acids in length.

Leptin, a protein hormone encoded by the *LEP* is secreted from white adipose tissue. This protein is found to regulate feed intake, energy partitioning, and metabolism (Liefers *et al.*, 2002; Lagonigro *et al.*, 2003), as well as lactogenesis (Feuermann *et al.*, 2004). The hypothalamus centrally mediates most of the effects of leptin. Leptin exerts its effect by stimulating or inhibiting the release of a neurotransmitter (neuropeptide Y) that eventually results in the decrease of feed intake and an increase in energy expenditure among other things (Houseknecht *et al.*, 1998).

Structure of the leptin protein

Leptin has a total pre-propeptide length of 167 amino acids, but the first 21 amino acids (the signal peptide) are excised, and the remaining 146 amino acids are released into the blood as a circulating protein. The binding of leptin to its receptor occur at the interface of two α-helices (Hiroike *et al.*, 2000).

In an analysis of the crystalline form of leptin using nuclear magnetic resonance, Zhang *et al.* (1997) demonstrated that leptin is a four-helix protein (A-B-C-D) whose structure is similar to that of the cytokine-family of proteins. The C and D helices are linked together by a single disulphide bond via a cysteine molecule on each of the helices, and this bond has been established to be crucial in maintaining the structure and stability of leptin (Rock *et al.*, 1996) (Figure 1.4). The binding of leptin to its receptor occur at the interface of the A and C α-helices (Hiroike *et al.*, 2000).

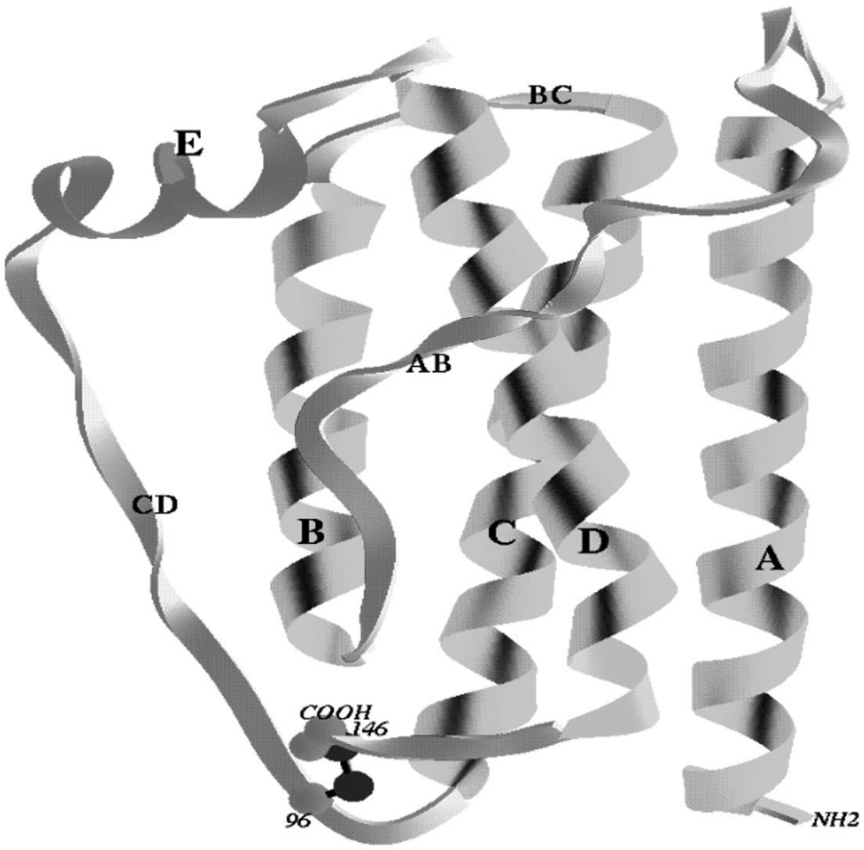


Figure 1.4 Structure of Leptin.
The leptin protein is made up of 146 amino acids, after the first 21 amino acids have been cleaved off. It has a four helical regions (A-B-C-D), similar to that of a cytokine four-helix bundle, and has a molecular weight of approximately 16kDa with a single disulphide bond that stabilises the protein, linking two cysteine molecules located at position 96 and 146 (joining helix C and D; Zhang *et al.*, 1997).

The regulatory function of leptin through the hypothalamus-hypophysis axis

The hypothalamus is identified as the main site of leptin's activity in regulating food intake and energy expenditure. Leptin receptor messenger RNA (mRNA) has been identified in the hypothalamus of several species, including humans (Considine *et al.*, 1996), rodents (Brogan *et al.*, 2000; Garcia *et al.*, 2000; Seeber *et al.*, 2002) and ruminants (Dyer *et al.*, 1997; Ren *et al.*, 2002). Leptin signals are converted into neural responses, and this results to changes in feed intake (Tang-Christensen *et al.*, 1999). A neurotransmitter identified as neuropeptide Y (NPY) is associated with regulation of food intake, hence leptin exerts its effect by stimulating or inhibiting the release of NPY, which eventually results in the decrease of feed intake and an increase in energy expenditure, among other things (Houseknecht *et al.*, 1998).

An increase in the levels of NPY mRNA were observed in mice deficient (*ob/ob*) or unresponsive (*db/db*) to leptin, but these changes were reversed in *ob/ob* mice following a direct administration of leptin (Ahima *et al.*, 2000). It was also observed that the production of hormones such as the gonadotrophins, lutenizing hormone (LH) and follicle stimulating hormone (FSH) from the hypophysis in the hypothalamus, were initiated following the administration of leptin (Woller *et al.*,

2001; Watanobe *et al.*, 2002; Amstalden *et al.*, 2003). Leptin also directly affects growth hormone-releasing hormone (GHRH)-mediated GH (growth hormone) secretion from the hypophysis (McMahon *et al.*, 2001; Zieba *et al.*, 2003a). Since Ob-R mRNA was identified in the anterior hypophysis of several species including sheep (Dyer *et al.*, 1997), pigs (Lin *et al.*, 2000), and mice and rats (Jin *et al.*, 2000), this could suggest that there may be some hypophysis effects of leptin on the secretion of FSH and LH as well.

Effect of leptin on food intake and bodyweight

The administration of leptin to animals has been revealed to produce a pronounced effect on food intake and bodyweight. Several reports have shown a drastic reduction in food intake and bodyweight for *ob/ob* mice (Campfield *et al.*, 1995), wild-type mice (Halaas *et al.*, 1995), monkeys (Tang-Christensen *et al.*, 1999) and pigs (Barb *et al.*, 1998), following the administration of leptin. In addition to leptin's effect on food intake via the activities of the NPY neurons in the hypothalamus, there are suggestions that leptin could also be regulating fat mobilization (Halaas *et al.*, 1995), as deduced from an investigation where *ob/ob* mice were pair-fed with leptin-treated *ob/ob* animals with the same amount of feed, yet, the *ob/ob* mice were found to lose 30% less weight compared to the leptin treated *ob/ob* mice (Halaas *et al.*, 1995).

In sheep, injection of recombinant human leptin into ewes for three days was found to decrease the dry matter intake (DMI) by 1/3 (Henry *et al.*, 1999). However, the aforementioned effect on DMI was lost after the sheep were starved and injected with leptin (Henry *et al.*, 2001; Morrison *et al.*, 2001). These changes to DMI are an indication that perhaps there is a different signal blocking leptin's effect on feed intake when the animal is in energy deficit. Furthermore, the level of plasma leptin in ruminants is affected by changes in feed accessibility. An increase in plasma leptin levels were observed in pregnant ewes and adult rams, within 48 hours respectively following an increase in feed intake (Blache *et al.*, 2000; Thomas *et al.*, 2001). In sheep, total deprivation to feed rapidly decreased plasma leptin levels (Marie *et al.*, 2001) and long-term restriction to feed decreased the concentration of plasma leptin (Delavaud *et al.*, 2000; Ehrhardt *et al.*, 2000; Morrison *et al.*, 2001). In cows, the concentration of leptin mRNA in adipose tissue was reduced when cows were deprived feed (Amstalden *et al.*, 2000).

Effect of leptin on body fatness (adiposity)

Adiposity, in the context of body fatness, has a positive correlation with the concentration of plasma leptin as has been observed in rodents and humans (Maffei *et al.*, 1995; Bunger *et al.*, 1999). In ruminants, either at the growing or lactating stage, it was observed that as adiposity increases, plasma leptin also increases and vice-versa (Blache *et al.*, 2000; Chilliard *et al.*, 2000; Thomas *et al.*, 2001). In a study involving dairy cows in late lactation, the level of plasma leptin was found to be positively correlated with body-condition score (BSC) (Ehrhardt *et al.*, 2000), although Delavaud *et al.* (2002) argued that the linear relationship is more to the size of fat cells, than to the BCS.

Leptin and the Mammary Gland

In ruminants, the expression of leptin mRNA in mammary tissue has been detected in both *in-vitro* and *in-vivo* investigations (Smith and Sheffield, 2002). During pregnancy and lactation in sheep, the two forms (long and short) of leptin receptor were expressed in the mammary gland. The concentrations of the two forms of leptin receptor peaked at the early- and mid-pregnancy phases, when active growth of the mammary gland is initiated, but decreased towards late pregnancy and the start of lactation (Bonnet *et al.*, 2002). Therefore, leptin and its receptor could be an important mediator of mammary gland growth and development (Maffei *et al.*, 1995; Laud *et al.*, 1999).

The effect of Leptin on lactation

The concentration of leptin is high during the early to mid-stage of pregnancy, but decreases markedly in the late stage of pregnancy towards parturition (Liefers *et al.*, 2003). In rats and cows, studies have shown that when the energetic cost of lactation is removed via the prevention of milk delivery, the concentration of leptin as well the energy balance, is increased (Woodside *et al.*, 2000; Block *et al.*, 2001). This suggests that the decrease in the concentration of leptin towards and during lactation is caused by the energy demand associated with milk production, especially since the suckling effect itself had no impact on the level of leptin (Brogan *et al.*, 1999).Taking into account the findings of Pickavance *et al.* (1998), whose report suggested that hypoleptinaemia may be critical in promoting the hyperphagia of lactation, it is logical to assume that the decrease in leptin level towards parturition is caused by the negative energy balance, whereas the low leptin levels during lactation is most likely responsible for the hyperphagia of lactation.

Leptin and Pregnancy

Studies have shown that the level of circulating leptin increased during early to mid-pregnancy and persists until late pregnancy in humans (Mukherjea *et al.*, 1999), rat (Amico *et al.*, 1998; Garcia *et al.*, 2000), mice (Tomimatsu *et al.*, 1997) and sheep (Ehrhardt *et al.*, 2001; Forhead *et al.*, 2002). This increase is as a result of increase in adiposity brought about by a corresponding increase in the expression of leptin mRNA in adipose tissue (Ehrhardt *et al.*, 2001). The increased level of leptin observed during pregnancy is unexpected, and this is because the time of pregnancy corresponds to a time of increased demand for nutrients and thus an increase in the activity of leptin is not expected. However, it is important to note that although the level of circulating leptin is increased during pregnancy, most of the leptin is in an inactive form (bound), which induces a change in its bioavailability (Mounzih *et al.*, 1998; Seeber *et al.*, 2002).

Nucleotide sequence variation in the bovine leptin gene

The bovine leptin gene is a highly polymorphic gene, and nucleotide sequence variation has been identified in both its coding and non-coding sequences. These include the presence of microsatellites and single nucleotide substitutions (Stone *et al.*, 1996b; Konfortov *et al.*, 1999). In an investigation of the exons and intronic regions of bovine *LEP* in a variety of cattle of *Bos taurus* and *Bos indicus*

origin, Konfortov *et al.* (1999), reported 20 nucleotide variations, six of which occurred in the exons, and at an approximate frequency of one per 84 bp. The remaining fourteen were observed in introns 1 and 2 with a slightly lower frequency of one per 92 bp, this giving an overall frequency of one per 89 bp. In Ensembl (EMBL-EBI, Hinxton, United Kingdom, release 96 - April 2019 ARS-UCD1.2) the cow leptin gene currently is described as containing 175 upstream variants, 3 variants in the 5' UTR region, 576 intron 1 variants, 5 exon 2 variants (both synonymous and non-synonymous), 97 intron 2 variants, 28 exon 3 variants (synonymous and non-synonymous), 255 variants in the 3'UTR and 242 variants downstream of the gene.

Several associations between variation in the leptin gene and traits of economic importance in many species have been reported. For example in humans, polymorphisms in the leptin gene have been associated with low circulating leptin levels (Hager *et al.*, 1998), birthweight variation (Orbak *et al.*, 2001) and obesity (Ohshiro *et al.*, 2000). In pigs, polymorphisms in the porcine leptin gene were associated with fatness (Jiang and Gibson, 1999) and in sheep, ovine leptin gene polymorphism was associated with weaning weight (Hajihosseinlo *et al.*, 2012). In cattle, several polymorphisms in the bovine leptin gene or leptin receptor gene have been described (Pomp *et al.*, 1997; Fitzsimmons *et al.*, 1998; Haegeman *et al.*, 2000) and associated with fat deposition in beef cattle (Fitzsimmons *et al.*, 1998), carcass FA composition (Kawaguchi *et al.*, 2017), milk yield (Banos *et al.*, 2008), milk fat levels (Giblin *et al.*, 2010; De Matteis *et al.*, 2012), milk protein levels (Giblin *et al.*, 2010), and milk FA composition (Pegolo *et al.*, 2016). A more detailed description of these nucleotide sequence variations is given in chapter 4.

In this thesis, the search for genetic variation in an extended region of bovine *MSTN* and *LEP* genes in a variety of cattle breeds from New Zealand and Nigeria will be undertaken using the Polymerase Chain Reaction (PCR) coupled with Single Strand Conformation Polymorphism (SSCP) analysis (PCR-SSCP). Nucleotide sequencing will then be employed to more precisely describe the nucleotide sequence differences observed. The PCR-SSCP technique is inexpensive and has the ability to screen for variation in a large number of cattle breeds, thus, giving a better representation of the entire breed. It is a reliable, reproducible and effective analytical method for the detection of deletions, insertions, or rearrangement in PCR-amplified DNA sequence (Hayashi, 1991; Konstantinos *et al.*, 2008).

1.4 Research on other genes and their association with milk production and composition of milk FAs traits

Several studies have described associations between genes and milk FA levels and profiles. For example, the diacylglycerol O-acyltransferase gene *(DGAT1)* has been associated with milk SFA yield (Winter *et al.*, 2002), the stearoyl-CoA desaturase gene (*SCD1)* has been associated with the UFA milk ratio (Garnsworthy *et al.*, 2010), and three haplotypes (H1, H2 and H3) of the FA binding protein

4 gene (*FABP4)* have been revealed to affect C14:0 levels in the early stage of lactation (Nafikov *et al.*, 2013).

However, these studies and many more were all based on the candidate genes approach, with focus on forward genetics. Therefore, there is need to investigate other genes whose protein products are expressed in the mammary gland and are therefore thought to affect milk production, as well as genes implicated with feed intake using the reverse genetics approach.

In this context, since the expression of MSTN occurs in the mammary gland of mouse (Manickam *et al.*, 2008) and pigs (Ji *et al.*, 1998), it suggests a possible role for MSTN in the development of the mammary gland and/or lactation. Given that the leptin gene (*LEP*) has been implicated with feed intake and lactogenesis (Liefers *et al.*, 2002; Feuermann *et al.*, 2004), investigating the effects of these two genes on milk production traits and milk composition is therefore important.

1.5 Aim and objectives of this study

The main aim of this study was to identify associations between variation in *MSTN* and *LEP* and gross milk traits and specific FA composition, and to further ascertain whether these would have potential value as gene-markers in improving the quality and quantity of bovine milk.

The first objective is to identify nucleotide sequence variation in an extended region of *MSTN* and *LEP* across a variety of cattle breeds from New Zealand and Nigeria, using polymerase chain reaction –single stranded conformation polymorphism (PCR-SSCP) analysis coupled with nucleotide sequencing.

The second objective is to attempt to define haplotypes spanning the amplified regions of the genes, provided a good number of homozygous and heterozygous samples are obtained.

The third objective is to examine whether variation in these genes, if found, is associated with variation in milk production traits and the component levels of milk FAs in NZ HF × J-cross cows, and if the genes are associated with any of the aforementioned traits, then ascertain whether gene-marker systems for these traits could be developed.

Chapter 2

Investigation of genetic variation and haplotypic diversity in the myostatin gene of New Zealand and Nigerian cattle breeds

With an increasing human population comes the need for improved agricultural technologies to cater for the growing global demand for food. In this context, the meat and milk industry would benefit from the development of improved ways of producing products with increased efficiency. Having a good understanding of the genes that underpin animal growth, carcass, milk production and other performance traits is therefore of importance.

One gene of recognised importance to animal growth, carcass and performance traits is the myostatin (MSTN) gene (*MSTN*), sometimes called the Growth and Differentiation Factor 8 (GDF8) gene (*GDF8*). Myostatin is a circulating factor, secreted by muscle cells, and whose function is to regulate the pre-natal proliferation of muscle fibres (McPherron *et al.*, 1997; Gonzalez-Cadavid *et al.*, 1998). Genetic variation in *MSTN* has been identified in cattle and linked to having increased numbers of muscle fibres, or what has been called double-muscling, in a number of breeds (Kambadur *et al.*, 1997; McPherron and Lee, 1997, Grobet *et al.*, 1997; Grisolia *et al.*, 2009).

There has been a huge documented evidence on genetic variations identified in *MSTN* across a variety of species. For example, in sheep, 78 nucleotide substitutions have been identified in various sheep breeds, some of which have resulted in a lack of biological function of the gene (Boman and Våge, 2009; Boman *et al.*, 2009; Boman *et al.*, 2010; Clop *et al.*, 2006; Gan *et al.*, 2008; Heaton *et al.*, 2007; Hickford *et al.*, 2010; Kijas *et al.*, 2007; Sjakste *et al.*, 2011; Zhou *et al.*, 2008). In mice, variations have been associated with a double-muscled phenotype as described by McPherron *et al.* (1997), and in cattle, nine variations have been linked with the double-muscled phenotypes (Dunner *et al.*, 2003; Esmailizadeh *et al.*, 2008; Grisolia *et al.*, 2009; Grobet *et al.*, 1998; Grobet *et al.*, 1997; Kambadur *et al.*, 1997). At least 20 different genetic variants (deletions, insertions and nucleotide substitutions) have been described in cattle *MSTN* (Aiello *et al.*, 2018).

While several studies have provided evidence of genetic variation in bovine *MSTN* and its effect on growth, carcass traits and performance traits, most of these studies have been focused on investigating specific gene regions in European breeds of cattle. For example, the seminal work of McPherron and Lee (1997) and Kambadur *et al.* (1997), reported on the effects of *MSTN* variation in Belgian Blue and Piedmontese cattle respectively. The Belgian Blue cattle *MSTN* had an 11-nucleotide deletion in the third exon that caused a frameshift mutation, and the Piedmontese cattle *MSTN* had a missense mutation in exon 3, which was also associated with increased muscle. In their study, Kambadur *et al.* (1997), reported that upon finding the 11-nucleotide deletion, further pedigrees of Belgian Blue cattle were tested for this deletion, including 16 pedigrees in New Zealand and four in the United States of

America (USA). Their results revealed that all the double-muscled purebred animals tested, were homozygous for this deletion.

Grobet *et al.* (1998) analysed ten European beef cattle breeds in which double-muscled animals have been described: including Belgian Blue cattle from Belgium; Blonde d'Aquitaine, Charolais, Gasconne, Limousin, Maine-Anjou, and Parthenaise cattle from France; Asturiana and Rubia Gallega cattle from Spain; and Italian Piedmontese cattle. They also analysed two Jersey cattle and two Holstein-Friesian cattle as 'controls'. Using a Polymerase Chain Reaction (PCR) and sequencing approach, they revealed seven DNA sequence variants within the coding region, including the previously described 11-nucleotide deletion, and four DNA sequence variants in intron sequences that they claimed were 'probably neutral'. Dunner *et al.* (2003) studied 678 cattle from 28 European breeds using a Single Strand Conformation Polymorphism (SSCP) method, defining seven new sequence variations, and concluding that *MSTN* was highly polymorphic having defined twenty haplotypes with approximately one nucleotide substitution every 100 bp.

Little effort has been made to characterise variation in *MSTN* in New Zealand cattle breeds, and while these breeds are predominantly derived from European/United Kingdom stock, little is known about how they relate to the current parent breeds. In the same manner, no studies on variation in *MSTN* have been conducted on the Nigerian cattle breeds. The Nigerian *Bos indicus* cattle belong to the Zebu cattle group and primarily comprise White Fulani (WF), Red Bororo (RB) and Sokoto Gudali (SG) breeds farmed for meat and milk purposes. They have a fatty thoracic hump on their shoulders and a large dewlap, and they are adapted to dry environmental conditions and high temperatures as well as being more resistant to tick infestation when compared to *Bos taurus* cattle (Mattioli *et al.*, 2000).

In this chapter, an investigation of genetic variation in bovine *MSTN* across a variety of common cattle breeds from New Zealand and Nigeria was undertaken. A PCR-SSCP analysis and DNA sequencing were used to investigate genetic variation in selected regions of *MSTN* in these cattle breeds, and subsequently an attempt was made in resolving the haplotypes across two polymorphic regions of *MSTN* in the cattle breeds investigated.

2.1 Materials and Methods

This research involving animals was performed in accordance with the Animal Welfare Act 1999 (New Zealand Government), whether blood was collected in New Zealand or Nigeria.

A total of 453 cattle was investigated, with these being of a variety of breeds found in New Zealand and Nigeria. These included the NZ Hereford (n = 30), Angus (n = 16), South Devon (n = 5), Composite (n = 3), Charolais (n = 9), Red Poll (n = 19), Shorthorn (n = 18), Simmental (n = 15), Murray Grey (n = 17), NZ HF x J-cross cattle (n = 160), and the Nigerian Sokoto Gudali (n = 3), Red

Bororo (n=13), White Fulani (n = 123) and cross-bred Holstein-Friesian × White Fulani cattle (n = 22). These cattle were genotyped in order to identify nucleotide sequence variation across an extended region of *MSTN*. Of these fourteen breeds, the NZ HF × J-cross cattle were bred for milk production, the Nigerian Sokoto Gudali, Red Bororo, White Fulani and cross-bred Holstein-Friesian × White Fulani were bred for milk and meat purposes, whereas the other nine New Zealand breeds are farmed primarily for meat production. The NZ Angus and Hereford are the preferred maternal breeds in New Zealand, in part because of their hardiness and suitability for farming in extensive hill and high country (rangeland) farming systems.

2.1.1 Blood collection and DNA extraction

Blood samples were collected from each animal onto FTA cards (Whatman, Middlesex, UK) by either piercing the ear of the animal or from the tail. The samples were air dried and DNA was purified using a two-step procedure as described by Zhou *et al.* (2006).

2.1.2 Polymerase Chain Reaction (PCR) amplification of targeted MSTN sequences

The polymerase chain reaction was used to amplify five different regions of *MSTN* (Figure 2.1). The reactions were performed in 15-µL volume containing the genomic DNA on a 1.2-mm diameter disc of FTA card, 0.25 µM for each primer, 150 µM for each dNTP (Eppendorf, Hamburg, Germany), 3.0 mM Mg^{2+}, 0.5 U of *Taq* DNA polymerase (Qiagen, Hilden, Germany), and 1× the reaction buffer supplied with the enzyme.

Amplification was undertaken in Bio-Rad S1000 thermal cyclers (Bio-Rad, Hercules, CA, USA). The thermal cycling conditions included an initial denaturation at 94 ºC for 2 minutes, followed by 35 cycles of 94 ºC for 30 seconds, annealing for 30 seconds at 58 ºC, or 60 ºC, depending on primer pairs (Figure 2.1), extension at 72 ºC for 30 seconds, and a final extension step at 72 ºC for 5 minutes.

The amplicons obtained from the PCR reactions were visualized by electrophoresis in 1% agarose (Quantum Scientific, Queensland, Australia) gels, using 1× TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM Na_2EDTA), containing 200 ng/mL ethidium bromide. A 2-µL aliquot of PCR product was added to 2 µL of loading dye (0.2% bromophenol blue, 0.2% xylene cyanol, 40% (w/v) sucrose) and the gels were run at a constant 10 V/cm for 10 minutes, prior to visualization by UV transillumination at 254 nm.

2.1.3 Detection of genetic variation using Single Stranded Conformation Polymorphism (SSCP) analyses

The amplicons obtained from the PCR reactions were subjected to SSCP analysis. A 0.7-µL aliquot of the amplicons was added to 7 µL of loading dye containing 10 mM EDTA, 0.025% bromophenol blue, 0.025% xylene-cyanol, and 98% formamide. The samples were then denatured at 95 ºC for 5 minutes, prior to snap chilling on wet ice. They were then loaded onto 16 cm x 18 cm, acrylamide: bisacrylamide (37.5:1) (Bio-Rad) gels. Electrophoresis was carried out using Protean II xi cells (Bio-

Rad) for 19 hours at various voltages, gel concentrations and room temperatures as described in Figure 2.1.

To detect the SSCP patterns, the gels were silver stained using a method described by Byun *et al.* (2009).

2.1.4 Detection of genetic variation using nucleotide sequencing

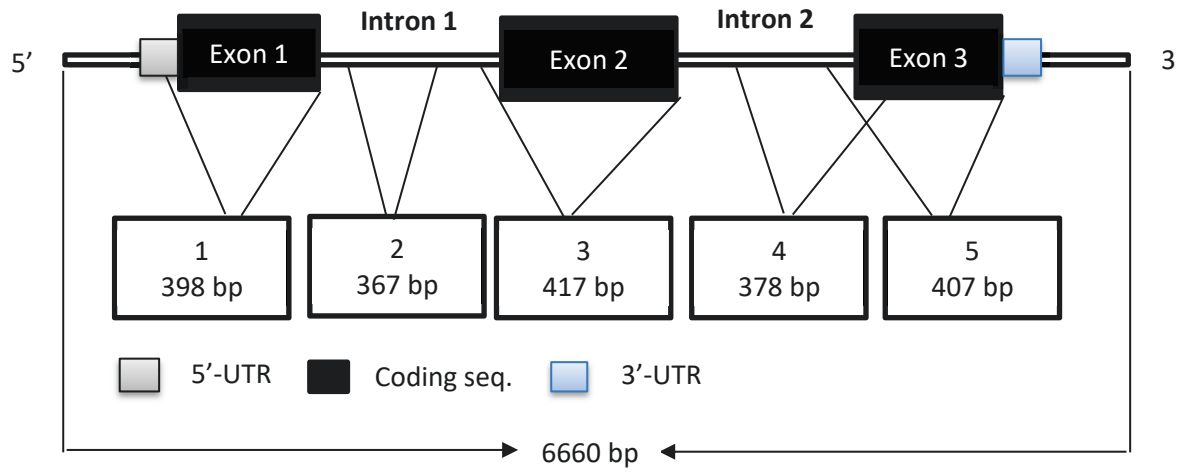
Based on the PCR-SSCP patterns observed, cattle that were homozygous with unique banding patterns, were subjected to direct sequencing. For heterozygous variants, the unique band was excised from the wet gel and incubated in water at 69 ºC for 1 hour. A 1-µL aliquot of the incubated product was aspirated and mixed with 14-µL of PCR pre-mixture for amplification via PCR reaction. Using a MiniElute™ PCR Purification Kit (Qiagen), DNA from the PCR products were purified and sequenced by the Lincoln University DNA Sequencing Facility, using the original PCR primers to prime the sequencing reactions. Each purified amplicon was sequenced in both directions, using the forward and reverse primers. The integrity of the sequence reactions was confirmed by viewing the peaks using Sequence Scanner v1.0 (Applied Biosystems). The sequence results were then subjected to subsequent sequence alignment and other comparisons using DNAMAN (Version 5.2.10, Lynnon BioSoft, Vaudreuil, Canada).

2.1.5 Determination of haplotypes

In determining the extended haplotypes of *MSTN*, two sets of primers were used to generate two amplicons (amplicons 2 and 4), and specific sequences were identified for each variant using the SSCP and sequencing methods described above. A combination of different PCR-SSCP pattern enabled the identification of the haplotypes using the following technique: cattle homozygous for one PCR-SSCP pattern in a region of the gene (e.g. *AA* for amplicon 2), were then genotyped in another region of the gene. If heterozygous in the second region (e.g. *AB* for amplicon 4), then two haplotypes (*A-A* and *A-B*), spanning the first and second regions could be defined.

2.2 Results

In this study, five pairs of primers were used to genotype 453 cattle from fourteen different breeds of New Zealand and Nigerian cattle. A total of five amplicons covering 1,967 bp of *MSTN* were generated. The amplified regions included exons 1, 2 and 3 and parts of introns 1 and 2 (Figure 2.1).



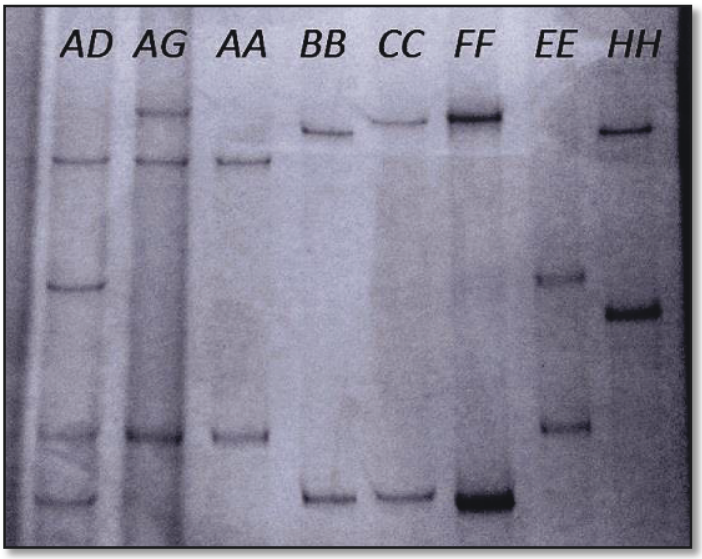
Amplicon	Primer sequence (5'-3')	Annealing temperature	Optimized SSCP conditions	Number of SSCP patterns
1	GGAAGAAGTAAGAACAAGGG CACCAGCAGGACTACTCAC	60 °C	12 % gel, 280 V, 13 °C	1
2	CATGGTACTATTGTTGAGAG AAGGCAAATCTATTCCAGG	58 °C	14 % gel, 390 V, 7 °C	8
3	GATATGGAGGTGTTTGTTCG CAGTAATCACTTACCAGTCC	58 °C	14 % gel, 390 V, 13 °C	1
4	ATAAGCAGGAAGACATAAGC GGTGTGTCTGTTACCTTGAC	58 °C	14 % gel, 390 V, 7 °C	3
5	CTCTTCTTTCCTTTCCATAC AAGACCTTCCATGTTTGAG	60 °C	11 % gel, 200 V, 20 °C	1

Figure 2.1 Primer pairs used and the amplicons generated from the PCR-SSCP analysis of bovine *MSTN*.
Primers were derived from bovine *MSTN* sequence with GenBank accession number AB076403. Optimized PCR-SSCP conditions for each amplicon and number of banding patterns are also shown.

2.2.1 PCR-SSCP patterns detected

Five different PCR-SSCP analyses coupled with DNA sequencing were used to search for genetic variation in amplicons 1 - 5 in all the cattle investigated. This technique revealed eight banding patterns (*A-H*) in amplicon 2 (within the intron 1 region) (Figure 2.2) and four patterns (A_4-D_4) in amplicon 4 (a region which encompasses the intron 2 - exon 3 boundary; Figure 2.3). No sequence variation was observed in the other amplified regions (Figure 2.4).

(a)



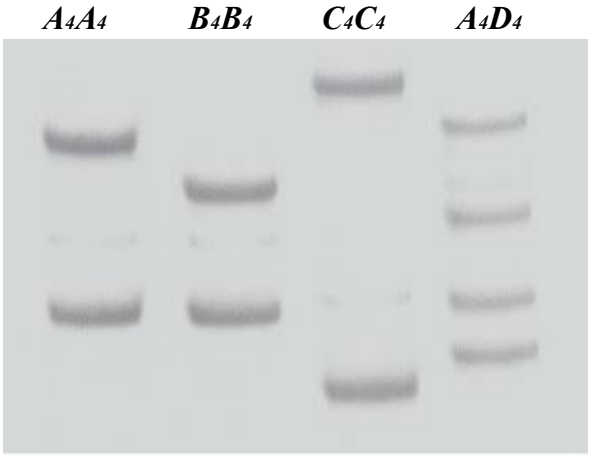
(b)

SNPs	*Variants*							
	A	***B***	***C***	***D***	***E***	***F***	***G***	***H***
c.373+751G/T	G	G	T	G	G	G	G	G
c.373+803T/G	T	T	G	G	G	G	G	T
c.373+877A/G	A	A	G	A	A	G	G	A
c.373+895G/C	G	G	C	G	G	G	G	G
c.374-909C/T	C	T	T	T	C	T	C	C
c.374-842G/C	G	G	G	C	C	G	G	G
c.374-812A/G	A	A	A	A	A	G	G	A
c.374-796C/T	C	C	C	C	C	C	C	T
c.374-762C/T	C	C	C	C	C	C	C	T

Figure 2.2 PCR-SSCP patterns detected for *MSTN* intron 1 region.
(a) Eight banding patterns (*A, B, C, D, E, F, G* and *H)* were detected in the 367 bp fragment of *MSTN* intron 1 investigated, seven (*A, B, C, D, E, F* and *G)* of which were observed in the New Zealand *Bos taurus* breeds, while four (*A, B, C and H*) were observed in Nigerian *Bos indicus* breeds. (b) Nine single nucleotide substitutions were detected in the eight variants.

Intron 2/exon 3 region

(a)



(b)

SNPs	Variants			
	A4	***B4***	***C4***	***D4***
c.748-78del1	-	-	-	T
c.748-195C/T	C	C	C	T
c.748-196C/T	C	C	C	T
c.748-202G/A	G	G	A	G
c.748-243G/C	G	G	C	G
c.748-281C/G	G	C	G	C
c.748-341T/C	T	T	C	T
c.748-350C/T	C	C	C	T
c.748-352C/T	T	C	C	C

Figure 2.3 PCR-SSCP banding patterns detected in *MSTN* intron 2 region

(a) Four banding patterns (*A_4, B_4, C_4* and *D_4)* were detected in the region of intron 2/exon 3 boundary investigated. The *A_4, B_4* and *D_4* variants were common in the New Zealand *Bos taurus* and Nigerian *Bos indicus* breeds, whereas the *C_4* variant was only found in the *Bos indicus* breeds. (b) Nucleotide sequencing revealed eight nucleotide substitutions and one deletion.

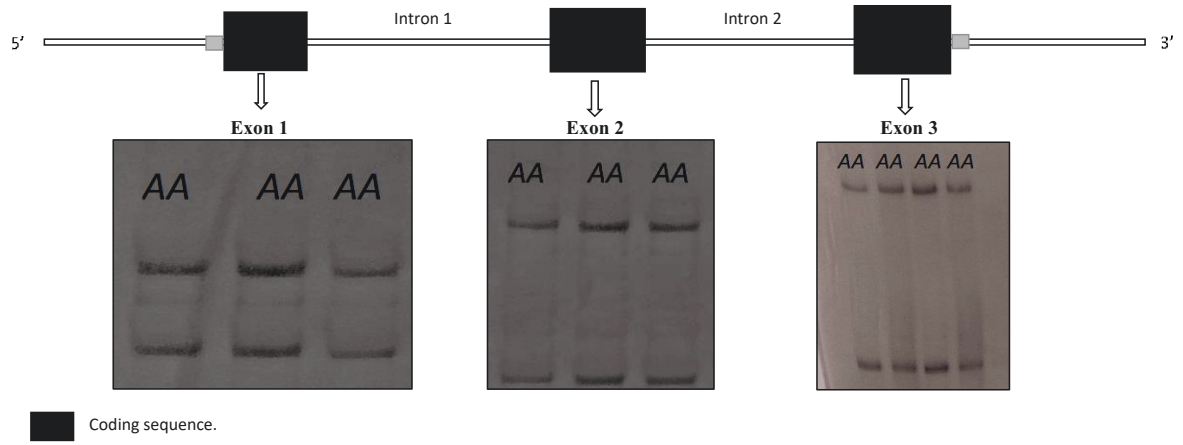


Figure 2.4 PCR-SSCP patterns of the coding sequences in bovine *MSTN*.
There was no genetic variation observed in exons 1, 2 and 3 of the bovine *MSTN* investigated in the New Zealand and Nigerian breeds.

Of the eight variants identified in the intron 1 region, the *A*, *B* and *C* variants were found in most of the New Zealand and Nigerian breeds investigated, and the *D*, *E*, *F* and *G* variants were only found in the New Zealand cattle breeds, whereas the *H* variant was observed only in the Nigerian breeds. The NZ Hereford and NZ HF x J-cross cows had five variants each (*A, B, C, D* and *E*) and (*A, B, C, D,* and *G*) respectively, while two variants (*A* and *B*) were observed in Simmental cattle, three variants (*A, B* and *C*) in Composite and Red Poll cattle, variants (*A, B* and *E*) in Angus and South Devon cattle, variants (*A, C* and *E*) in Charolais cattle, and four variants (*A, B, C* and *F*) and (*A, B, C* and *E*) in Shorthorn and Murray Grey cattle respectively. The Red Bororo, White Fulani and cross-bred Holstein-Friesian × White Fulani breeds had variants *A, B, C* and *H*, while the Sokoto Gudali had the *A, B* and *H* variants. In the intron 2 - exon 3 region, all the fourteen breeds had the A_4, B_4 and D_4 variants except the Red Poll, Composite and Sokoto Gudali breeds, which carried only the A_4 and B_4 variants. The C_4 variant was only detected in all the Nigerian breeds except the Sokoto Gudali. A summary table of breeds and the respective variants are presented in table 2.1.

Table 2.1 The different variants found in each of the fourteen cattle breeds investigated.

Breeds	Intron 1 Variants*								Intron 2 Variants			
	A	***B***	***C***	***D***	***E***	***F***	***G***	***H***	A_4	B_4	C_4	D_4
Hereford	*	*	*	*	*	-	-	-	*	*	-	*
Angus	*	*	-	-	*	-	-	-	*	*	-	*
South Devon	*	*	-	-	*	-	-	-	*	*	-	*
Charolais	*	-	*	-	*	-	-	-	*	*	-	*
Red Poll	*	*	*	-	-	-	-	-	*	*	-	-
Shorthorn	*	*	*	-	-	*	-	-	*	*	-	*
Simmental	*	*	-	-	-	-	-	-	*	*	-	*
Composite	*	*	*	-	-	-	-	-	*	*	-	-
Murray Grey	*	*	*	-	*	-	-	-	*	*	-	*
Holstein-Friesian × Jersey cross-bred	*	*	*	*	-	-	*	-	*	*	-	*
Sokoto Gudali	*	*	-	-	-	-	-	*	*	*	-	-
Red Bororo	*	*	*	-	-	-	-	*	*	*	*	*
White Fulani	*	*	*	-	-	-	-	*	*	*	*	*
Cross bred White Fulani × Holstein Fresian	*	*	*	-	-	-	-	*	*	*	*	*

*The symbol (*) represents the presence of a variant in a breed, while (-) represents the absence of a variant.

2.2.2 Nucleotide variations detected

A total of 18 nucleotide variations were identified across the regions examined. Seventeen were nucleotide substitutions (c.373+751G/T, c.373+803T/G, c.373+877A/G, c.373+895G/C, c.374-909C/T, c.374-842G/C, c.374-812A/G, c.374-796C/T and c.374-762C/T located in intron 1, and c.748-195C/T, c.748-196C/T, c.748-202G/A , c.748-243G/C, c.748-281C/G, c.748-341T/C , c.748-350C/T, c.748-352C/T located in the intron 2 - exon 3 boundary region. One of the nucleotide variations was a deletion c.748-78del1 (Table 2.2). Of the seventeen substitutions, seven (c.373+751G/T, c.373+803T/G, c.373+877A/G, c.373+895G/C, c.374-796C/T, c.374-762C/T and c.748-350C/T) were novel and are reported here for the first time. The remaining ten, with their respective rs numberings (c.374-909C/T rs109597350, c.374-842G/C rs110496111, c.374+812A/G rs385362581, c.748-195C/T rs207514521, c.748-196C/T rs210984981, c.748-202G/A (rs477807563), c.748-243G/C (rs459203170) c.748-281C/G rs110910793, c.748-341T/C(rs483297434) and c.748-352C/T rs132859209) had been previously reported to Ensembl (EMBL-EBI, Hinxton, United Kingdom, release 93 - July 2018 UMD3.1:GK000002.2 for cow MSTN). The deletion c.748-78del1 was identified at the intron 2 - exon 3 boundary has been described previously by Grobet *et al.* (1998).

Table 2.2 Eighteen nucleotide variations identified in bovine *MSTN* across fourteen different New Zealand and Nigerian cattle breeds and crosses

	SNPs																	
	Intron 1									Intron 2								
[#]BREEDS	c.373+751	c.373+803	c.373+877	c.373+895	c.374-909	c.374-842	c.374-812	c.374-796	c.374-762	[&]c.748-78del	c.748-195	c.748-196	c.748-202	c.748-243	c.748-281	c.748-341	c.748-350	c.748-352
Hereford	G/T	T/G	A/G	G/C	C/T	G/C	A	C	C	-/T	C/T	C/T	G	G	C/G	T	C/T	C/T
Angus	G	T/G	A	G	C/T	G/C	A	C	C	-/T	C/T	C/T	G	G	C/G	T	C/T	C/T
S. Devon	G	T/G	A	G	C/T	G/C	A	C	C	-/T	C/T	C/T	G	G	C/G	T	C/T	C/T
Charolais	G/T	T/G	A/G	G/C	C/T	G/C	A	C	C	-/T	C/T	C/T	G	G	C/G	T	C/T	C/T
Red Poll	G/T	T/G	A/G	G/C	C/T	G	A	C	C	-	C	C	G	G	C/G	T	C	C/T
Shorthorn	G/T	T/G	A/G	G/C	C/T	G	A/G	C	C	-/T	C/T	C/T	G	G	C/G	T	C/T	C/T
Simmental	G	T	A	G	C/T	G	A	C	C	-/T	C/T	C/T	G	G	C/G	T	C/T	C/T
Composite	G/T	T/G	A/G	G/C	C/T	G	A	C	C	-	C	C	G	G	C/G	T	C	C/T
M. Grey	G/T	T/G	A/G	G/C	C/T	G/C	A	C	C	-/T	C/T	C/T	G	G	C/G	T	C/T	C/T
NZ HF × J	G/T	T/G	A/G	G/C	C/T	G/C	A/G	C	C	-/T	C/T	C/T	G	G	C/G	T	C/T	C/T
S. Gudali	G	T	A	G	C/T	G	A	C	C	-/T	C/T	C/T	G	G	C/G	T	C/T	C/T
R. Bororo	G/T	T/G	A/G	G/C	C/T	G	A	C/T	C/T	-/T	C/T	C/T	G/A	G/C	C/G	T/C	C/T	C/T
W. Fulani	G/T	T/G	A/G	G/C	C/T	G	A	C/T	C/T	-/T	C/T	C/T	G/A	G/C	C/G	T/C	C/T	C/T
HF × WF	G/T	T/G	A/G	G/C	C/T	G	A	C/T	C/T	-/T	C/T	C/T	G/A	G/C	C/G	T/C	C/T	C/T

[&]The symbol (-) represents a deletion

[#] S. Devon=South Devon; M. Grey=Murray Grey; NZ HF × J= New Zealand Holstein-Friesian × Jersey-cross; S. Gudali=Sokoto Gudali; R. Bororo=Red Bororo; W. Fulani=White Fulani; HF × WF = Holstein-Friesian × White Fulani cross.

The numbering is based on the recommended nucleotide nomenclature (http://www.hgvs.org/mutnomen/recs-DNA.html#number).

2.2.3 Haplotypes resolved across MSTN polymorphic regions

A total of 20 haplotypes (H1-H20) were unambiguously defined spanning amplicon 2 (intron 1) and amplicon 4 (the intron 2 - exon 3 boundary region) in the fourteen breeds investigated. Some of these haplotypes were exclusive to a breed, while others appeared to be common across the breeds. For example, the haplotypes H1-H8 were found in all the New Zealand and Nigerian breeds investigated, whereas the haplotypes H19 and H20 were only found in the Nigerian breeds. Haplotype H19 carrying the c.748-341C, c.748-281C, c.748-243C, c.748-202A, c.748-78del1 was observed in all Nigerian breeds except the Sokoto Gudali, while H20, carrying c.374-796T, c.374-762T, c.748-352C, c.748-281C, c.748-78del1 was observed in all the Nigerian breeds. The haplotype H11, carried SNPs c.373+803G, c.374-909T, c.374-842C and c.748-78del1, and was only observed in the Hereford cattle. Similarly, haplotype H14 with SNPs c.373+803G, c.374-842C, c.748-350T, c.748-196T and c.748-195T, was observed only in the South Devon cattle. Haplotypes H12, H16, H17 and H18 were exclusive to the NZ HF x J-cross cattle and haplotype H2, containing the deletion c.748-78del1 was the most common, and was found in all breeds, except the South Devon cattle (Table 2.3).

Table 2.3 The frequency of *MSTN* haplotypes identified in New Zealand and Nigerian cattle breeds.

		[#]Haplotypes Frequency %																		
[$]BREEDS	**n**	**H1**	**H2**	**H3**	**H4**	**H5**	**H6**	**H7**	**H8**	**H9**	**H10**	**H11**	**H12**	**H13**	**H14**	**H15**	**H16**	**H17**	**H18**	**H19**	**H20**
Hereford	28	7.14	21.43	-	-	3.57	7.14	17.86	3.57	-	21.43	3.57	-	14.30	-	-	-	-	-	-	-
Angus	9	-	50.00	6.25	-	12.50	18.80	-	-	-	-	-	-	12.50	-	-	-	-	-	-	-
South Devon	3	-	-	-	-	-	50.00	-	-	-	-	-	-	-	50.00	-	-	-	-	-	-
Charolais	9	12.50	75.00	-	-	-	-	-	6.25	-	-	-	-	6.25	-	-	-	-	-	-	-
Red Poll	12	25.00	33.33	-	16.67	16.70	-	8.33	-	-	-	-	-	-	-	-	-	-	-	-	-
Shorthorn	17	-	86.35	-	-	-	-	-	4.55	4.55	-	-	-	-	-	4.55	-	-	-	-	-
Simmental	14	8.30	4.20	-	20.80	12.50	54.20	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Murray Grey	17	4.55	72.70	-	-	-	-	9.10	9.10	4.55	-	-	-	-	-	-	-	-	-	-	-
HF × J- cross	36	9.20	11.11	3.70	11.11	9.20	5.50	18.47	7.50	-	1.90	-	3.70	-	-	-	1.90	11.11	5.60	-	-
Sokoto Gudali	3	25.00	25.00	-	25.00	25.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Red Bororo	13	62.50	-	-	6.25	-	-	6.25	-	-	-	-	-	-	-	-	-	-	-	18.75	6.25
White Fulani	98	51.11	8.16	-	3.06	-	1.02	8.16	-	-	-	-	-	-	-	-	-	-	-	19.34	9.15
HF × WF cross	46	34.50	5.17	1.72	10.35	-	1.72	15.52	3.45	-	-	-	-	-	-	-	-	-	-	6.90	20.67

[$]The Composite cattle are excluded from this table, as a consequence of the absence of cattle with a homozygous genotype that would enable haplotype determination.

[#] Symbol (-) signifies the absence of that haplotype within the breed.

Table 2.4 Nucleotides present for the SNPs identified in the 20 haplotypes (H1-H20) defined across the two regions of *MSTN* investigated.

	Intron 1								SNPs	Intron 2							
[#]Haplotype	c.373+751	c.373+803	c.373+877	c.373+895	c.374-909	c.374-842	c.374-812	c.374-796	c.374-762	c.748-78del1	c.748-195	c.748-196	c.748-202	c.748-243	c.748-281	c.748-341	c.748-350	c.748-352
H1	G	T	A	G	C	G	A	C	C	-	C	C	G	G	G	T	C	T
H2	G	T	A	G	C	G	A	C	C	-	C	C	G	G	C	T	C	C
H3	G	T	A	G	C	G	A	C	C	T	T	T	G	G	C	T	T	C
H4	G	T	A	G	T	G	A	C	C	-	C	C	G	G	G	T	C	T
H5	G	T	A	G	T	G	A	C	C	-	C	C	G	G	C	T	C	C
H6	G	T	A	G	T	G	A	C	C	T	T	T	G	G	C	T	T	C
H7	T	G	G	C	T	G	A	C	C	-	C	C	G	G	G	T	C	T
H8	T	G	G	C	T	G	A	C	C	-	C	C	G	G	C	T	C	C
H9	T	G	G	C	T	G	A	C	C	T	T	T	G	G	C	T	T	C
H10	G	G	A	G	T	C	A	C	C	-	C	C	G	G	G	T	C	T
H11	G	G	A	G	T	C	A	C	C	-	C	C	G	G	C	T	C	C
H12	G	G	A	G	T	C	A	C	C	T	T	T	G	G	C	T	T	C
H13	G	G	A	G	C	C	A	C	C	-	C	C	G	G	C	T	C	C
H14	G	G	A	G	C	C	A	C	C	T	T	T	G	G	C	T	T	C
H15	G	G	G	G	T	G	G	C	C	T	T	T	G	G	C	T	T	C
H16	G	G	G	G	C	G	G	G	C	-	C	C	G	G	G	T	C	T
H17	G	G	G	G	C	G	G	G	C	-	C	C	G	G	C	T	C	C
H18	G	G	G	G	C	G	G	G	C	T	T	T	G	G	C	T	T	C
H19	G	T	A	G	C	G	A	C	C	-	C	C	A	C	C	C	C	C
H20	G	T	A	G	C	G	A	T	T	-	C	C	G	G	C	T	C	C

[#]The symbol (-) represents a deletion and the highlighted nucleotides indicate the variation.

2.3 Discussion

With a view to having a comprehensive understanding of natural genetic variation occurring at the bovine *MSTN* locus, this study focused on investigating the genetic diversity in fourteen different breeds of New Zealand *Bos taurus* and Nigerian *Bos indicus* cattle that are used for meat and/or milk production. A PCR-SSCP approach and DNA sequencing were used to investigate an extended region of the gene that spanned nucleotide c.-25 in the 5'-UTR to c.*24 in the 3'-UTR. This encompassed the three exons, and parts of introns 1 and 2. Using this approach, previously described sequence variations were identified, as well as some novel variations that are reported here for the first time.

Of the 18 nucleotide variations identified across the regions examined, one was a deletion and seventeen were nucleotide substitutions. The deletion was located at position c.748-78del1 near the intron 2 - exon 3 boundary. This deletion is very close to the splice-donor site, and it has been previously identified in European cattle breeds (Grobet *et al.*, 1998). It is becoming more evident that introns are functionally active participants in gene and genome functionality. This is because they can encode regulatory elements (Yutzey *et al.*, 1989) that participate in splicing, transcription, and recombination events. This means that both conserved and variable non-coding sequences may possess functionality, especially since the variable region could also be adaptively active. Details on how variation in or around the splice donor cite can affect gene expression and/or protein assembly as illustrated by Sjakste *et al.* (2011) is discussed in chapter four. In effect, sequence variation at, or near intron/exon boundaries may affect gene expression and to confirm this in the context of our investigation, further studies of MSTN expression in cattle carrying the c.748-78del1 deletion would be required.

Of the remaining seventeen nucleotide substitutions, ten (c.374-909C/T rs109597350, c.374-842G/C rs110496111, c.374+812A/G rs385362581, c.748-195C/T rs207514521, c.748-196C/T rs210984981, c.748-202G/A rs477807563, c.748-243G/C rs459203170, c.748-281C/G rs110910793, c.748-341T/C rs483297434 and c.748-352C/T rs132859209) have been described at Ensembl. Of these ten substitutions, two (c.748-281C/G and c.748-352C/T) have been reported previously (He *et al.*, 2013), in a study involving Chinese and Red Angus cattle. They examined an extended region of the *MSTN*, which spanned from the promoter to the 3'-UTR in 16 cattle of the Qinchuan breed and four of the Red Angus breed. Upon comparing sequences from the Qinchuan cattle with a *Bos taurus* sequence (GenBank accession number AF320998), a total of 69 SNPs were identified in the Qinchuan breed. Of these 69 SNPs, one was a 16-bp insertion observed in the intron 1 of three Qinchuan cattle, and this has also been previously reported (Tantia *et al.*, 2006).

In the intron 1 region investigated in this study, variants *A, B* and *C* with accession numbers (MK353501, MK353502 and MK353503) respectively, carrying nucleotide variations; c.373+751T, c.373+803G, c.373+877G, c.373+895C, and c.374-909T were found in both the New Zealand *Bos*

taurus and the Nigerian *Bos indicus* breeds. Likewise, in the intron 2 region, variants A_4, B_4 and D_4 with accession numbers (MK353508, MK353509 and MK353510) carrying SNP's; c.748-78del, c.748-195T, c.748-196T, c.748-281G, c.748-350T, and c.748-352T were also observed in the New Zealand and Nigerian breeds.

Edwards *et al.* (2007), had suggested that *Bos primigenius* is the progenitor of all taurine and zebu (African *Bos indicus*) cattle. In this context, the identification of common variants across the New Zealand cattle of *Bos taurus* origin and the Nigerian cattle of Zebu origin investigated here, may suggest an evolutionary relationship between these two species. This is supported by the findings of Hiendleder *et al.* (2008), whose report suggested an estimate of divergence time for *Bos taurus* and *Bos indicus* from a common ancestor to be pre-Neolithic. More details about the relationship between *Bos taurus* and *Bos indicus* are discussed in chapter four.

However, the *H* variant carrying SNPs c.374-796T and c.374-762T in the intron 1 region, and the C_4 variant carrying SNPs c.748-202A, c.748-243C and c.748-341C in the intron 2 region, were only found in the Nigerian *Bos indicus* samples. The identification of these variants only in the Nigerian breeds might be a function of species differences and/or number of animals investigated. Either way, what is interesting and worthy of note is the fact that, the Nigerian *Bos indicus* breeds had unique variants *H* and C_4 in the intron 1 and intron 2 regions respectively. Therefore, it is very likely that these unique variants are species-specific, and as such, could suggest that both *H* and C_4 variants are unique to the Nigerian *Bos indicus* species. It'll be interesting to see what phenotypic trait these variants may be influencing. Such variations occurring between species are to be expected, partly because they form the basis for species differences, especially if they occur either in the coding or non-coding sequences, since it is variation within these sequences that produce phenotypic variation between both individuals and species (Mattic, 2001).

In the intron 1 region investigated, nine SNPs were identified. He *et al.* (2013), reported four SNPs (c.373+807G/T, c.373+881A/G, c.373+899G/A, and c.373+934C/T) that are different to the ones identified here, in this same region. With a potential total of at least thirteen SNPs clustered in this region, it appears this part of intron 1 in *MSTN* is highly variable and thus, it may be under less functional constraint. This has been suggested by Hickford *et al.* (2010), in a study of ovine *MSTN* intron 1, which involved 747 sheep from a variety of New Zealand breeds and breed composites. Their study revealed a total of seven SNPs in a region of intron 1 of ovine *MSTN*, four of which were clustered in the region c.373 + 240 to c.373 + 249, this suggesting that this part of the intron is subject to less functional constraint.

Since these variations in sheep and cattle occur in a non-expressed region of *MSTN*, it is perhaps less obvious how they might affect the structure or function of the protein. They might however influence the rate of transcription and/or translation of the gene, as was described by Liu *et al.* (1995) in an investigation involving both transgenic mice and *in vitro* studies, where intron sequences in the gene

encoding rat growth hormone stimulate transcription by promoting assembly of an ordered nucleosome array in the vicinity of the promoter. Additionally, He *et al.* (2010) transformed C2C12 cell-lines with a transgene construct that contained part of the bovine *MSTN* promoter (pMD-MSTNPro) and another construct containing the first intron of bovine *MSTN* (pMD-Intron1). They observed an increase in the mean fluorescence intensity of Green Fluorescent Protein (GFP) gene and percentage of fluorescence positive cells and concluded that the presence of intron 1 of bovine *MSTN* improved the expression of GFP in the transformed cells. Conversely, while there is little or no information on the possible effect of the absence of bovine *MSTN* intron 1 on gene expression, it would be interesting to explore this path, especially, since most of the previous studies (McPherron *et al.*, 1997) have only focused on examining the effects of 'knocking-out' the active portion of the MSTN peptide, and not the introns.

There was no genetic variation identified in the coding sequences of bovine *MSTN* in this study. This is consistent with the studies of Clop *et al.* (2006) and Kijas *et al.* (2007). However, Dunner *et al.* (2003) identified five nucleotide variations in the coding region of bovine *MSTN,* two of which brought about an amino acid change, while the others were silent. These included p.S105C, which was described in Parthenaise cattle and p.D182N resulting from a G/A substitution in the Maine-Anjou breed. A silent exon 1 substitution c.267A/G was found in Aubrac, Bazadaise and Salers cattle, whereas c.324C/T in exon 1 was observed in Maine-Anjou, Charolais, Aubrac, and Salers cattle, and an Intr95 sire-line. A third silent substitution, c.387G/A, was found in exon 2 in Ayrshire, Maine-Anjou, Salers, and Galloway cattle.

A possible explanation for why exon variation was not found in both the New Zealand and Nigerian cattle may be because of the breeds investigated. While Dunner *et al.* (2003), investigated a total of 28 breeds, only nine: the Parthenaise, Main-Anjou, Salers, Aubrac, Bazadaise, Galloway, INTR95 sire-line, Ayrshire and Charolais were reported to have variation in the coding regions of the gene. Of these nine breeds, only the Charolais was among the breeds investigated in this study, and the Charolais cattle investigated did not have the c.324C/T substitution. This could have been for a number of reasons including test sample size, founder effects, selection, including cross-breeding and in-breeding. Some of these reasons have been suggested previously by Blott *et al.* (1998), who revealed genetic differences between Hereford cattle from Britain, New Zealand, Canada, Ireland and Sweden. In a study by Kantanen *et al.* (2000b), it was suggested that genetic divergence in Nordic cattle breeds and within population diversity, can be explained by the combined effects of breed origin, admixture during foundation and development and random genetic drift due to limitations in effective population size; either when the breed was founded, or more recently. Therefore, while most of the New Zealand breeds might have originated from Europe, a number of different effects may have contributed to differences existing between these breeds and their European ancestors.

Some of the variants were found to be common to more than one breed, while others were apparently rare and breed-specific. For example, of the eight variants defined in intron 1, variant *A* was found in all the fourteen breeds and crosses. The Hereford cattle and the NZ HF x J-cross cows shared variants *A, B, C* and *D*, while the Composite, Red Poll, Murray Grey and Shorthorn cattle shared the *A, B* and *C* variants. In the intron 2 - exon 3 region studied, variants A_4, B_4 and D_4 were found in all the breeds except the Red Poll and Composite cattle, which carried only the A_4 and B_4 variants. Of the nine New Zealand 'beef breeds' investigated, five (Hereford, Angus, South Devon, Murray Grey and Charolais) shared intron 1 *E*. These five breeds are among the most common beef breeds in NZ, hence sharing *E* may be indicative of among other things having a common ancestor, or from having been historically inter-bred, or because of selection for important meat/carcass-related traits. In the latter context, *E* might be associated with improved performance, although the absence of *E* in the other beef cattle (Composite cattle, Red Poll, Shorthorn and Simmental) might weaken this selection argument, and overall care needs to be taken in the context of only a small number of any of the breeds (except the NZ HF x Jersey-cross dairy cows), having been studied. Equally, these cattle were not, and probably could not have been selected so as to be representative of the whole breed population in New Zealand, as all the cattle studied were in private ownership.

While SNPs and localised sequence variation may provide some indication of the structural diversity of a gene and uniqueness of a breed or breeds, haplotypes spanning a region are typically more informative, especially if they encompass all or most of the coding sequences. A total of 20 haplotypes (H1 to H20) spanning intron 1 (amplicon 2) to the intron 2 - exon 3 boundary region (amplicon 4) of *MSTN* were resolved (Table 2.3). It has been shown that SNPs that are adjacent or close to each other tend to be co-inherited and thus usually show strong linkage disequilibrium (Hey, 2004). This phenomenon was observed with the *MSTN* haplotypes, as they could be separated into seven broad haplotype groups based on the location of the SNPs. For example, with the intron 1 region, haplotypes H1 to H6 carried c.374-909T (named Group 1), while H7 to H9 with SNP's c.373+751T, c.373+803G c.373+877G, c.373+ 895C and c.374-909T cluster as Group 2, whereas H10 to H14 with SNP's c.373+803G, c.374-909T and c.374-842C cluster as Group 3, and H15 to H18 with SNPs c.373+803G c.373+ 877G, c.374-909T and c.374-812G cluster as Group 4. In the intron 2 region, haplotypes H3, H6, H9, H12, H14, H15 and H18, which carry c.748-195T, c.748-196T and c.748-350T could be clustered as Group 5, H1, H4, H7, H10 and H16 with SNPs c.748-78del, c.748-281G and c.748-352T could be clustered as Group 6, and finally H2, H5, H8, H11, H13 and H17 with SNP c.748-78del could be clustered as Group 7.

Recombination activities do not occur randomly throughout the genome, but instead tend to occur in recombination hotspots (Jeffreys *et al.*, 2005), which are usually small regions characterized by significantly higher recombination rate than in surrounding regions. These hotspots are associated with trinucleotide repeats such as CGG-CCG, GAG-CTG, GAA-TTC, and GCN-NGC and are conserved in mammals and yeast (Aguilera *et al.*, 2008). Investigations have suggested that these hotspots tend to

form hairpin structures on the lagging strand during replication from single-stranded DNA base-pairing with itself in the trinucleotide repeat region (Aguilera *et al.*, 2008). These hairpin structures cause DNA breaks that lead to a higher frequency of recombination at these sites (Aguilera *et al.*, 2008). In this investigation, some of these trinucleotides associated with DNA hotspots in the regions of introns 1 and 2 investigated were observed. Some occurred as a single sequence, while others occurred as a repeat in the positions where the nucleotide variations were identified. For example, the trinucleotide 'CTG' was found in the intron 1 position where SNP c.373+751G/T was detected. Also, the trinucleotide sequence 'GAA-TTC' was observed in the region of intron 2 where the adjacent SNPs (c.748-195C/T and c.748-196C/T) were identified, but it was only the trinucleotide 'TTC' that seemingly appeared as a repeated sequence in the positions where SNPs (c.748-350C/T, c.748-352C/T and c.748-78delT) were discovered. These trinucleotides were not observed as precisely repeated sequences in the regions where all the variations were found except for the 'TTC', hence it is difficult to conclude whether these regions of introns 1 and 2 are actual hotspots for DNA recombination.

However, it is important to note that the identification of these trinucleotide sequence(s) in the exact positions where the proximal SNP's were detected may not have been by chance. This could suggest that perhaps these trinucleotide sequences are capable of influencing DNA recombination irrespective of whether they appear as a repeated sequence, or as a single sequence. Further analysis of these regions of introns 1 and 2 will have to be undertaken to ascertain the possibilities of these regions serving as hotspots for DNA recombination.

The haplotype H2, carrying the deletion c.748-78del, occurred at a relatively high frequency in all the breeds except the South Devon cattle. The sharing of common haplotypes suggests that cattle might be more closely related, although this is not proof. Dunner *et al.* (2003) suggested that the pattern of haplotype sharing is an indicator of the history of the different bovine breeds, thus the distribution of shared haplotypes is useful in investigating population relationships. The presence of breed-specific haplotypes on the other hand may explain the differences between breeds, especially if the haplotypes are found within the coding sequences. In a study around the identification of shared haplotypes that are identical by descent (IBD), Ying *et al.* (2015), suggested that while common variants are mostly shared across ethnic groups, rare variants are more likely to be recent in history and population-specific. In this context, the occurrence of variants *F* and *G* exclusively in the Shorthorn and NZ HF x J-cross cows, suggests that these variants are breed-specific. Such variations occurring within a species might be expected, as they form the basis for individual differences within a population, especially because it is variation within both coding and non-coding sequences that produces phenotypic variation between both individuals in a species and between different species (Mattic, 2001). It will therefore be interesting to see what unique trait these variants may be influencing in the respective breeds and whether their presence could explain the differences between these breeds.

The haplotype H15 found in the Shorthorn cattle, and H16, H17 and H18 observed in the NZ HF x J-cross cows, share some rare but also more common variations (Table 2.3). This is not surprising, especially considering the evolution and introduction of the Shorthorn breed into New Zealand, where it was initially used for milking purpose (Te Ara-the Encyclopaedia of NZ, accessed 16 November, 2018), just like the pure and cross-bred Holstein-Friesian and Jersey cattle of today. Since these two breeds were farmed in New Zealand for the same purpose at some point in time, it is possible that they may have shared a common ancestor, and as such it may not be that surprising that they share some *MSTN* variations in common.

In the region of intron 1 investigated across the fourteen breeds and crosses, it is noteworthy that some breeds appear to have a reduced number of variants. For example, while the Hereford and the NZ HF x J-cross cows had five variants each, the Simmental had two variants, the Composite, Red Poll, Angus and South Devon had three variants each, and the Shorthorn and Murray Grey cattle were found to have four variants each. In the intron 2 - exon 3 region, all the breeds had three variants, except for the Red Poll cattle and the Composite cattle, which had two variants each. It is difficult to confirm whether this is a characteristic of the specific breeds, or a function of sampling and sample size. Interestingly, this phenomenon was also observed in sheep, as reported by Hickford *et al.* (2010) in a study involving 747 sheep in New Zealand, where only three of the five variants were observed in the NZ Romney sheep breed investigated. They suggested this could either be a unique characteristic of the breed, or perhaps the sheep investigated may not have been a good representation of the breed across New Zealand, either by number or by the method of selection.

2.4 Conclusion

Based on the seven novel substitutions identified in this chapter, together with the other ten substitutions and one deletion reported previously, it appears that *MSTN* is more highly variable than reported previously. It might also therefore be expected that as more cattle breeds are investigated, more variation(s) unique to those breeds, or perhaps uncommon in other breeds, might be identified. Overall, this study provides more insight into the conserved and polymorphic nature of the coding and non-coding sequences of bovine *MSTN* respectively. It also highlights possible linkages of bovine *MSTN* haplotypes based on proximal occurrence of SNPs. Given the potential of this gene to affect muscle growth and thus meat production, the findings lay a foundation to proceed to see if this genetic variation might be of value in improving production traits.

Chapter 3

Associations between bovine *MSTN* variants with milk production traits and fatty acids composition in New Zealand Holstein-Friesian × Jersey (HF × J)-cross cows

The myostatin gene (*MSTN*), alternatively known as the growth and differentiation factor 8 gene (*GDF8*), encodes the myostatin protein (MSTN). It has pleiotropic effects, and its expression has been associated with increased skeletal muscle mass, and decreased adipogenesis as a result of the reduced secretion of leptin (McPherron *et al.*, 1997; Whittemore *et al.*, 2003; Mendias *et al.*, 2008).

Sequence variation in *MSTN* has been associated with increases in growth and muscling traits in several species. For example, in cattle it has been linked to having increased numbers of muscle fibres, or double-muscling, and in a number of breeds (Kambadur *et al.*, 1997; McPherron and Lee, 1997). Similarly, sequence variation in the first intron has been found to influence growth and carcass traits such as leg, loin and total yield of lean meat in NZ Romney sheep (Hickford *et al.*, 2010).

While there is well documented evidence describing how variation in *MSTN* is associated with growth and carcass traits in cattle, there is little information around the effects of *MSTN* variations on milk yield and milk FA composition. There are suggestions that MSTN could affect lactation by affecting the production of FA in milk, and primarily through MSTN deficiency being associated with decreased adipogenesis (McPherron and Lee, 2002; Whittemore *et al.*, 2003; Mendias *et al.*, 2008).

Bovine milk fat is approximately 70% saturated fatty acid (SFA), 25% monounsaturated fatty acid (MUFA) and 5% polyunsaturated (PUFA) (Grummer, 1991). From a human health perspective, an increase in unsaturated fatty acid (UFA) content and a decrease in SFA content could be considered favourable; and equally from a physical perspective (e.g., increased spread-ability of butter), increased relative levels of UFA would be desirable. Previously, it has been identified that the concentrations of FAs in milk fat are affected most by four parameters: the diet of the cow (Palmquist *et al.*, 1993; Dewhurst *et al.*, 2006), genetic variation between cows within a breed (Soyeurt *et al.*, 2006b; Stoop *et al.*, 2008), breed differences (Palladino *et al*,. 2010; Maurice-Van Eijndhoven *et al.*, 2011) and the number of days in milk (DIM) (Strucken *et al.*, (2015). In their report, Strucken *et al.* (2015), established that cows in early lactation stage are characterized by negative energy balance as the dietary energy intake is unable to meet the demands of high milk production in approximately the first 60 days of lactation. To offset this balance, an alternative energy source is needed. This leads to the mobilisation of body energy stores to balance the deficit between feed intake, and energy expenditure on maintenance and milk production (Bauman *et al.*, 1980). However, as a consequence of the cow's body fat being used up for this purpose, other biological pathways are affected, resulting in a change in milk composition.

The Holstein-Friesian and the Jersey breeds have one of the most notable differences in terms of the composition of milk FAs. Milk from Jersey cows tends to have higher concentrations of some short- and medium-chain SFA, but lower concentrations of some UFA (Arnould and Soyeurt, 2009). These differences could be capitalized upon in order to obtain the preferred FA profile through the crossing of these breeds.

In the above context, this chapter will explore associations between *MSTN* variations, and key milk production traits and milk FA composition in NZ HF × J-cross dairy cows.

3.1 Materials and Methods

3.1.1 Cattle and milk sample collection

This research was approved by the Lincoln University Animal Ethics Committee under the provisions of the Animal Welfare Act 1999 (New Zealand Government).

A total of 430 NZ HF × J-cross dairy cows, of variable and unknown breed proportion, and of 3 to 13 years of age were studied. The cows were from two herds, and all the cows were grazed on pasture (a mixture of perennial ryegrass and white clover) on the Lincoln University Dairy Farm (LUDF; Canterbury, New Zealand). All the cows calved over the months of August-September and were milked twice a day.

Milk samples for gross milk trait analysis were collected once a month from September to February and the daily milk yield in litres was recorded using Tru-test milk meters (Tru-test Ltd, Auckland, New Zealand) and Fourier-Transform Infra-Red Spectroscopy (MilkoScan FT 120 Foss, Hillerød, Denmark) was used to analyse milk samples for fat percentage (%) and protein percentage (%). Average daily milk yield, and average protein and fat percentages were calculated over the 6 months of milk collection. The milk samples for FA analysis were collected at 148 ± 19 days in milk from each cow in a single afternoon milking in mid-January (the middle of summer in the Southern Hemisphere). These were frozen at -20 ^{o}C, and then freeze-dried, prior to being individually ground to a fine powder for component analysis.

3.1.2 Gas chromatography of the FAs in the milk samples

The FAs were methylated and extracted in n-heptane, before being analysed by Gas Chromatography (GC) as FA methyl esters (FAMEs). The methylation reactions were performed in 10-mL Kimax tubes. Individual freeze-dried and powdered milk samples (0.17 g), were dissolved in 900 µL of n-heptane (100%, AR grade), before 100 µL of internal standard (5 mg/ml of C21:0 methyl ester in n-heptane) and 4.0 mL of 0.5 M NaOH (in 100% anhydrous methanol) were added.

The tubes were vortexed then incubated in a block heater (Ratek Instruments, Australia) at 50 ℃ for 15 minutes. After cooling to room temperature, another 2.0 mL of n-heptane and 2.0 mL of deionized water was added to each tube. After vortexing, the tubes were centrifuged (Megafuge 1.0R, Heraeus, Germany) for 5 minutes at 1500*g*. The top layer of n-heptane was transferred into a second Kimax tube and 2.0 mL of n-heptane was added to each of the original tubes. The extraction was repeated and the n-heptane aspirates were then pooled. Anhydrous sodium sulphate (10 mg) was added to the n-heptane extracts, to remove any residual water.

The GC analysis was carried out using a Shimadzu GC-2010 Gas Chromatograph (Shimadzu Corporation, Kyoto, Japan) equipped with a flame ionization detector and an AOC-20i auto sampler. The output was analysed with GC Solution Software (Shimadzu). For analysis, 1 µL of the n-heptane sample extract was injected into a 100 metre GC column (250 µm × 0.25 µm capillary column, CP-Select, Varian) with a 1:60 split ratio. The separation was undertaken with a helium carrier gas and was run for 92 minutes. The temperature of both the injector and detector were set at 250 ℃ and the thermal profile of the column consisted of 45 ℃ for 4 minutes, followed by 27 minutes at 175 ℃ (ramped at 13 ℃/minute), 35 minutes at 215 ℃ (ramped at 4 ℃/minute), and a final 'bake-off' at 250 ℃ for 5 minutes (ramped at 25 ℃/minute.). The individual FAMEs were identified by the peak retention time compared to commercially obtained external standards (ME61, ME93, BR3, BR2, ME100, GLC411 and GLC463; Laroden AB, Sweden). Quantification of the individual FAMEs was based on peak area assessment and comparison with the internal and external standards. The threshold for peak area determination on the chromatogram was a 500-unit count, with peaks that were under 500-unit count, being ignored. The calculated minimum component of an individual FAME was therefore 0.01 grams per 100 grams of total FA.

After their individual measurement, the FAs were arranged into various groups and indices. These groups were, short chain-length FAs (SCFA) = C4:0 + C6:0 + C8:0; medium chain-length FAs (MCFA) = C10:0 + C12:0 + C14:0; long chain-length FAs (LCFA) = C15:0 + C16:0 + C17:0 + C18:0 + C19:0 + C20:0 + C22:0 + C24:0; omega 3 FAs = C18:3 *cis*-9, 12, 15 + C20:5 *cis*-5,8, 11, 14, 17 + C22:5 *cis*-7, 10, 13, 16, 19; omega 6 FAs = C18:2 *cis*-9, 12 + C18:3 *cis*-6, 9, 12 + C20:3 *cis*-8, 11, 14 + C20:4 cis-5, 8, 11, 14; monounsaturated FAs (MUFA) = C10:1 + C12:1 + C14:1 *cis*-9 + C15:1 + C16:1 *cis*-9 + C17:1 + C18:1 *trans*-11 + C18:1 *cis*-9 + C18:1 *cis*-(10 to 15) + C20:1 *cis*-5 + C20:1 *cis*-9 + C20:1 *cis*-11 + C22:1 *trans*-13; polyunsaturated FAs (PUFA) = C18:2 *trans*-9, 12 + C18:2 *cis*-9,*trans*-13 + C18:2 *cis*-9,*trans*-12 + C18:2 *trans*-9,*cis*-12 + C18:2 *cis*-9, 12 + C18:3 *cis*-6, 9, 12 + C18:3 *cis*-9, 12, 15 + CLA + C20:3 *cis*-8, 11, 14 + C20:4 *cis*-5, 8, 11, 14 + C20:5 *cis*-5, 8, 11, 14, 17 + C22:5 *cis*-7, 10, 13, 16, 19; and total branched FA= C13:0 *iso* + C13:0 *anteiso* + C15:0 *iso* + C15:0 *anteiso* + C17:0 *iso*.

Unsaturated FA indices were also calculated as follows: C12:1 index (C12:1 divided by the sum of C12:0 and C12:1); C16:1 index (C16:1 *cis*-9 divided by the sum of C16:0 and C16:1 *cis*-9) and C18:1

index (C18:1 *cis*-9 divided by the sum of C18:0 and C18:1 *cis*-9). The method is as described by Li *et al.* (2019), with the un-adjusted mean levels in the 430 cows being calculated and used subsequently in the statistical analyses.

3.1.3 Blood sample collection

Blood samples were collected from each cow onto FTA cards (Whatman, Middlesex, UK) by piercing the ear of the animal. The samples were air dried and DNA purification was carried out using a two-step procedure as described by Zhou *et al.* (2006).

3.1.4 Polymerase chain reaction amplification of MSTN

The intron 1 region of MSTN was amplified using PCR, with the forward and reverse primers (5'-catggtactattgttgagag-3' and 5'-aaggcaaatctattccagg-3' respectively) being the same as those used previously (Chapter 2). The PCR reaction was performed in 15-µL reactions containing the genomic DNA on a 1.2-mm diameter disc of FTA card, 0.25 µm for each primer, 150 µm for each dNTP (Eppendorf, Hamburg, Germany), 3.0 mM Mg^{2+}, 0.5 U of *Taq* DNA polymerase (Qiagen, Hilden, Germany*),* and 1 x the reaction buffer supplied with the enzyme.

Amplification was undertaken in Bio-Rad S1000 thermal cyclers (Bio-Rad, Hercules, CA, USA). The thermal cycling conditions included an initial denaturation at 94 ºC for 2 minutes, followed by 35 cycles of 94 ºC for 30 seconds, annealing for 30 seconds at 58 ºC, extension at 72 ºC for 30 seconds, and a final extension step at 72 ºC for 5 minutes.

3.1.5 Single Stranded Conformation Polymorphism (SSCP) analysis of MSTN

The SSCP technique was used to detect genetic variation in the amplicons obtained from the PCR reactions. A 0.7-µL aliquot of the amplicons was added to 7 µL of loading dye containing 10 mM ethylenediaminetetraacetic acid (EDTA), 0.025% bromophenol blue, 0.025% xylene-cyanol, and 98% formamide. The samples were then placed on a hot plate already set at 95 ºC for 5 minutes for denaturation, followed by snap chilling on wet ice. Samples were then loaded onto 16 cm × 18 cm, 14% acrylamide: bisacrylamide (37.5:1) (Bio-Rad) gels. Electrophoresis was carried out using Protean II xi cells (Bio-Rad) for 19 hours at 390 volts and 7 ºC in 0.5 × Tris/Borate/EDTA buffer.

To detect the SSCP banding patterns, the gels were silver stained using the method described by Byun *et al.* (2009).

3.1.6 Nucleotide sequencing

Based on the PCR-SSCP patterns observed, cattle that were homozygous with unique banding patterns were subjected to direct sequencing. For heterozygous variants, the unique band was excised from the wet gel and incubated in water at 69 ºC for 1 hour. A 1-µL aliquot of the incubated product was aspirated and mixed with 14 µL of PCR pre-mixture for re-amplification, and subsequently sequenced

as described by Gong *et al.* (2011). The sequences were aligned, and other analyses undertaken using DNAMAN (Version 5.2.10, Lynnon BioSoft, Vaudreuil, Canada).

3.1.7 Statistical analysis

Using an online chi-square calculator (http://www.oege.org/software/hwe-mr-calc.shtml), Hardy-Weinberg equilibrium (HWE) for the *MSTN* genotypes was analysed. The statistics software IBM SPSS version 22 (IBM, NY, USA) was used to perform all other statistical analyses, and unless otherwise stated significance was accepted when $P < 0.05$.

General Linear Mixed-effects Models (GLMMs) was used to test the associations between *MSTN* variation and variation in milk component levels. First, single-variant presence/absence models (fixed effects: days in milk (DIM), age, herd and *MSTN* variants) were used to ascertain which variant(s) should be analysed in subsequent multi-variant models. The multi-variant models included any variant that had a variant-FA trait association in the single-variant presence/absence analysis with a *P*-value of less than 0.200. This is a low threshold for the inclusion of an explanatory factor in the model. The multi-variant models were again corrected for the fixed effects of DIM, age, herd and *MSTN* variant, and with DIM and age fitted as covariates.

Next a GLMM (fixed effect: *MSTN* genotypes, DIM, age and herd) and multiple pair-wise comparisons with Bonferroni corrections were used to ascertain the effect of genotypes on milk FA component levels.

Unfortunately, the effect of sire of cow could not be included in the GLMMs, because some semen straws (sire genetics) used in New Zealand dairy cattle artificial insemination-based breeding approaches contain mixed-sire semen purchased from commercial semen producers. In these cases, it is impossible to ascertain individual sire identity. However, since the straws were mixed-semen straws and because different sires are used for different inseminations, in different years, it is unlikely that sire was a strongly confounding effect. Cow age and herd might also be confounded with sire, but this cannot be confirmed.

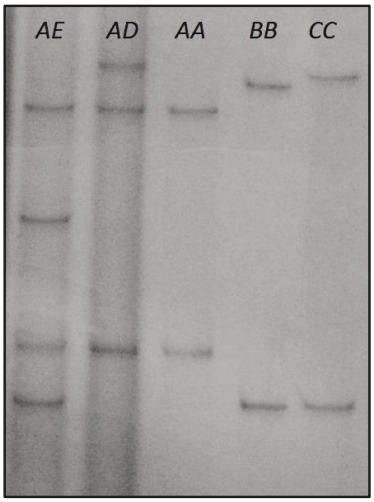
3.2 Results

3.2.1 Identification of nucleotide sequence variation in bovine MSTN

In this study, a pair of primer (5'-catggtactattgttgagag-3’ and 5’-aaggcaaatctattccagg-3’) was used to genotype 430 NZ HF × J-cross cows for a region of *MSTN*. A 367 bp fragment of the intron 1 region of bovine *MSTN* was amplified and analysed using PCR-SSCP analyses. These analyses coupled with DNA sequencing revealed five banding patterns (*A*-*E*) in the region of intron 1 investigated (Figure 3.1). A total of seven single-nucleotide substitutions (c.373+751G/T, c.373+803T/G, c.373+877A/G,

c.373+895G/C, c.374-909C/T, c.374-842G/C, and c.374-812A/G) were identified, all of which were described in Chapter 2.

(a)



(b)

		Variants				
Nucleotide position	**Variation**	***A***	***B***	***C***	***D***	***E***
c.373+751	G/T	G	G	T	G	G
c.373+803	T/G	T	T	G	G	G
c.373+877	A/G	A	A	G	G	A
c.373+895	G/C	G	G	C	G	G
c.374-909	C/T	C	T	T	C	T
c.374-842	G/C	G	G	G	G	C
c.374-812	A/G	A	A	A	G	A

Figure 3.1 Variation in intron 1 of bovine *MSTN* identified using PCR-SSCP (a) and confirmed by nucleotide sequencing (b). Five PCR-SSCP patterns, representing the five unique sequence variants (*A-E*), in both homozygous and heterozygous genotypes are shown.

3.2.2 Single-variant presence/absence models

General linear mixed-effect models revealed that the presence (or absence) of *MSTN B, C* and *D* in a cow's genotype was associated with the adjusted mean quantity of some individual and grouped milk FAMEs. The different variants had varying effects on the composition of milk FA levels as detailed in tables 3.1 and 3.2. Variant *D* was associated with reduced levels of C8:0, C10:0 and C12:0, grouped MCFA level, and the level of the MUFA C12:1 ($P < 0.05$), but increased levels of C16:1 *cis*-9 ($P < 0.05$). Variant *B* was associated with lower adjusted mean level of C15:0 *iso* ($P < 0.05$), whereas the presence of *C* was associated with higher C20:3 *cis*-8, 11, 14 and C22:1, *trans*-13 levels ($P < 0.05$).

The effect of variant *A* (the most common variant), and variant *C* on gross milk production traits were not significant (Table 3.3), but an interesting trend is still evident as *A* was showing a trend of association with higher total percentage of fat ($P = 0.064$) and protein ($P = 0.079$), while *C* decreased total percentage of fat ($P = 0.068$) and protein ($P = 0.084$).

3.2.3 Genotype analyses

The genotypes *AA* (n = 151), *AB* (n = 92), *AC* (n = 53) and *AD* (n = 65) occurred at a frequency greater than 5% and were analysed in this model. The other genotypes *AE* (n = 15), *BB* (n = 7), *BC* (n = 7), *BD* (n = 17), *CC* (n = 6) *CD* (n = 10) and *DD* (n = 7) had frequencies of less than 5% each, and were not included in the model (Figure 3.2).

When one copy of the *AA* (the most common genotype) is replaced by a *D* variant, the resulting *MSTN* genotype was found to have an effect on the composition of milk FA levels, which was consistent with the results of the single variant model. Genotype *AD* was associated with lower adjusted mean levels of C10:0, C12:0 and C12:1 (Table 3.4).

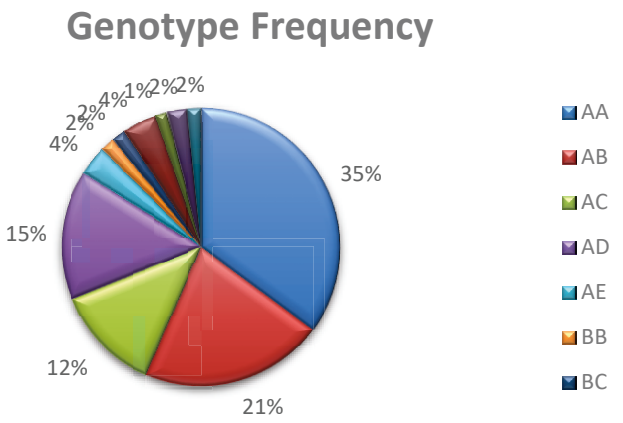


Figure 3.2 Pie chart showing genotypes and their respective frequencies obtained from the region of intron 1 in bovine *MSTN*. Only the genotypes *AA*, *AB*, *AC* and *AD* occurred at a frequency greater than 5% and were analysed in the statistical model.

Table 3.1 Associations between bovine *MSTN* variants and individual milk FAs in New NZ HF × J-cross cows

Individual Fatty Acids	Variant	Other variants in model	Mean FAME level ± SE[1] (g/100 g milk FA) Absent	n	Present	n	*P*
C4:0	*A*	none	1.280 ± 0.022	53	1.284 ± 0.017	377	0.898
	B	none	1.282 ± 0.017	302	1.270 ± 0.018	128	0.475
	C	none	1.283 ± 0.017	357	1.281 ± 0.021	73	0.999
	D	none	1.282 ± 0.017	337	1.291 ± 0.020	93	0.441
C6:0	*A*	none	1.551 ± 0.019	53	1.574 ± 0.014	377	0.479
	B	none	1.570 ± 0.015	302	1.560 ± 0.015	128	0.396
	C	none	1.561 ± 0.014	357	1.563 ± 0.018	73	0.722
	D	none	1.573 ± 0.014	337	1.551 ± 0.017	93	0.294
C8:0	*A*	none	1.175 ± 0.016	53	1.188 ± 0.013	377	0.343
	B	none	1.119 ± 0.013	302	1.176 ± 0.014	128	*0.164*
	C	none	1.185 ± 0.013	357	1.184 ± 0.015	73	0.917
	D	none	1.191 ± 0.012	337	1.165 ± 0.015	93	**0.021**
	B	*D*	1.185 ± 0.018	302	1.168 ± 0.019	128	*0.093*
	D	*B*	1.190 ± 0.014	337	1.163 ± 0.016	93	**0.014**
C10:0	*A*	none	3.226 ± 0.064	53	3.254 ± 0.050	377	0.605
	B	none	3.262 ± 0.051	302	3.223 ± 0.054	128	0.320
	C	none	3.244 ± 0.050	357	3.256 ± 0.061	73	0.807
	D	none	3.275 ± 0.049	337	3.161 ± 0.058	93	**0.009**
C10:1	*A*	none	0.274 ± 0.008	53	0.286 ± 0.006	377	*0.081*
	B	none	0.286 ± 0.006	302	0.278 ± 0.006	128	*0.066*
	C	none	0.284 ± 0.006	357	0.280 ± 0.007	73	0.439
	D	none	0.285 ± 0.006	337	0.278 ± 0.007	93	*0.187*
	A	*B, D*	0.275 ± 0.008	53	0.285 ± 0.007	377	*0.145*
	B	*A, D*	0.285 ± 0.007	302	0.276 ± 0.007	128	*0.050*
	D	*A, B*	0.284 ± 0.007	337	0.276 ± 0.008	93	*0.129*
C11:0	*A*	none	0.055 ± 0.003	53	0.058 ± 0.002	377	0.258
	B	none	0.058 ± 0.002	302	0.055 ± 0.003	128	*0.129*
	C	none	0.058 ± 0.002	357	0.054 ± 0.003	73	*0.082*
	D	none	0.058 ± 0.002	337	0.055 ± 0.003	93	0.211
	B	*C*	0.057 ± 0.003	302	0.054 ± 0.004	128	*0.073*
	C	*B*	0.058 ± 0.003	357	0.053 ± 0.003	73	**0.049**
C12:0	*A*	none	3.901 ± 0.085	53	3.953 ± 0.067	377	0.479
	B	none	3.964 ± 0.068	302	3.902 ± 0.073	128	0.233
	C	none	3.938 ± 0.067	357	3.948 ± 0.082	73	0.879
	D	none	3.980 ± 0.066	337	3.819 ± 0.078	93	**0.006**
C13:0 *iso*	*A*	none	0.077 ± 0.003	53	0.081 ± 0.002	377	*0.104*
	B	none	0.081 ± 0.002	302	0.078 ± 0.002	128	*0.097*
	C	none	0.080 ± 0.002	357	0.078 ± 0.003	73	0.404
	D	none	0.080 ± 0.002	337	0.077 ± 0.003	93	*0.090*
	A	*B, D*	0.077 ± 0.003	53	0.080 ± 0.003	377	0.279
	B	*A, D*	0.080 ± 0.003	302	0.077 ± 0.003	128	*0.063*
	D	*A, B*	0.080 ± 0.003	337	0.077 ± 0.003	93	*0.058*

(Continue on next page)

(Continued)

C12:1	*A*	none	0.089 ± 0.003	53	0.092 ± 0.002	377	*0.184*
	B	none	0.093 ± 0.003	302	0.089 ± 0.003	128	*0.050*
	C	none	0.092 ± 0.002	357	0.091 ± 0.003	73	0.845
	D	none	0.093 ± 0.002	337	0.088 ± 0.003	93	**0.022**
	A	*B, D*	0.090 ± 0.004	53	0.090 ± 0.004	377	0.870
	B	*A, D*	0.092 ± 0.004	302	0.088 ± 0.004	128	**0.025**
	D	*A, B*	0.093 ± 0.003	337	0.087 ± 0.004	93	**0.012**
C13:0 *anteiso*	*A*	none	0.037 ± 0.001	53	0.037 ± 0.001	377	0.763
	B	none	0.037 ± 0.001	302	0.037 ± 0.001	128	*0.201*
	C	none	0.037 ± 0.001	357	0.037 ± 0.001	73	0.714
	D	none	0.037 ± 0.001	337	0.037 ± 0.001	93	0.623
C13:0	*A*	none	0.113 ± 0.005	53	0.118 ± 0.004	377	0.245
	B	none	0.118 ± 0.004	302	0.114 ± 0.004	128	*0.172*
	C	none	0.118 ± 0.004	357	0.113 ± 0.004	73	*0.188*
	D	none	0.118 ± 0.004	337	0.114 ± 0.004	93	*0.198*
	B	*C, D*	0.116 ± 0.005	302	0.111 ± 0.005	128	*0.086*
	C	*B, D*	0.117 ± 0.005	357	0.111 ± 0.005	73	*0.093*
	D	*B, C*	0.116 ± 0.005	337	0.111 ± 0.005	93	*0.107*
C14:0	*A*	none	12.515 ± 0.148	53	12.482 ± 0.117	377	0.795
	B	none	12.505 ± 0.118	302	12.464 ± 0.126	128	0.652
	C	none	12.451 ± 0.116	357	12.607 ± 0.141	73	*0.170*
	D	none	12.528 ± 0.115	337	12.372 ± 0.136	93	*0.123*
	C	*D*	12.437 ± 0.126	357	12.583 ± 0.151	73	0.201
	D	*C*	12.536 ± 0.121	337	12.384 ± 0.142	93	*0.136*
C14:1 *cis*-9	*A*	none	0.940 ± 0.037	53	0.975 ± 0.029	377	0.263
	B	none	0.978 ± 0.030	302	0.950 ± 0.032	128	0.227
	C	none	0.971 ± 0.029	357	0.955 ± 0.035	73	0.590
	D	none	0.971 ± 0.029	337	0.954 ± 0.034	93	0.502
C15:0 *iso*	*A*	none	0.287 ± 0.005	53	0.291 ± 0.004	377	0.333
	B	none	0.292 ± 0.004	302	0.286 ± 0.004	128	**0.031**
	C	none	0.288 ± 0.004	357	0.294 ± 0.004	73	*0.148*
	D	none	0.291 ± 0.004	337	0.286 ± 0.004	93	*0.087*
	B	*C, D*	0.291 ± 0.005	302	0.284 ± 0.005	128	**0.018**
	C	*B, D*	0.287 ± 0.005	357	0.291 ± 0.006	73	0.347
	D	*B, C*	0.291 ± 0.005	337	0.284 ± 0.005	93	**0.045**
C15:1	*A*	none	0.275 ± 0.005	53	0.279 ± 0.004	377	0.346
	B	none	0.280 ± 0.004	302	0.275 ± 0.005	128	*0.282*
	C	none	0.277 ± 0.004	357	0.281 ± 0.005	73	0.300
	D	none	0.278 ± 0.004	337	0.277 ± 0.005	93	0.650
C15:0 *anteiso*	*A*	none	0.617 ± 0.016	53	0.626 ± 0.013	377	0.537
	B	none	0.628 ± 0.013	302	0.617 ± 0.014	128	0.273
	C	none	0.621 ± 0.013	357	0.632 ± 0.015	73	0.361
	D	none	0.626 ± 0.013	337	0.616 ± 0.015	93	0.374
C15:0	*A*	none	1.430 ± 0.030	53	1.456 ± 0.024	377	0.306
	B	none	1.456 ± 0.024	302	1.440 ± 0.025	128	0.376
	C	none	1.450 ± 0.023	357	1.448 ± 0.029	73	0.910
	D	none	1.455 ± 0.023	337	1.432 ± 0.028	93	0.263
C16:0	*A*	none	36.940 ± 0.518	53	37.175 ± 0.408	377	0.593
	B	none	37.014 ± 0.412	302	37.290 ± 0.441	128	0.388
	C	none	37.259 ± 0.505	357	36.694 ± 0.493	73	*0.254*
	D	none	37.038 ± 0.404	337	37.370 ± 0.476	93	0.350

(Continue on next page)

(Continued)

C16:1 *cis*-9	*A*	none	1.358 ± 0.043	53	1.353 ± 0.034	377	0.894
	B	none	1.361 ± 0.034	302	1.343 ± 0.037	128	0.506
	C	none	1.355 ± 0.034	357	1.349 ± 0.041	73	0.853
	D	none	1.338 ± 0.034	337	1.402 ± 0.040	93	**0.031**
C17:0 *iso*	*A*	none	0.562 ± 0.012	53	0.559 ± 0.009	377	0.719
	B	none	0.560 ± 0.009	302	0.558 ± 0.010	128	0.788
	C	none	0.557 ± 0.009	357	0.567 ± 0.011	73	0.253
	D	none	0.561 ± 0.009	337	0.556 ± 0.011	93	0.555
C17:0	*A*	none	0.702 ± 0.012	53	0.701 ± 0.683	377	0.920
	B	none	0.698 ± 0.009	302	0.706 ± 0.010	128	0.264
	C	none	0.700 ± 0.009	357	0.703 ± 0.011	73	0.736
	D	none	0.703 ± 0.009	337	0.695 ± 0.011	93	0.339
C17:1	*A*	none	0.212 ± 0.004	53	0.208 ± 0.003	377	0.327
	B	none	0.209 ± 0.003	302	0.208 ± 0.004	128	0.674
	C	none	0.208 ± 0.003	357	0.212 ± 0.004	73	0.281
	D	none	0.208 ± 0.003	337	0.213 ± 0.004	93	*0.300*
C18:0	*A*	none	8.823 ± 0.220	53	8.582 ± 0.173	377	*0.298*
	B	none	8.599 ± 0.175	302	8.706 ± 0.187	128	0.429
	C	none	8.611 ± 0.173	357	8.728 ± 0.210	73	0.486
	D	none	8.636 ± 0.172	337	8.652 ± 0.203	93	0.911
C18:1 *cis*-9	*A*	none	13.756 ± 0.263	53	13.447 ± 0.207	377	*0.167*
	B	none	13.509 ± 0.210	302	13.539 ± 0.224	128	0.850
	C	none	13.482 ± 0.206	357	13.636 ± 0.251	73	0.445
	D	none	13.452 ± 0.205	337	13.730 ± 0.242	93	*0.124*
	A	*D*	13.755 ± 0.270	53	13.470 ± 0.219	377	0.208
	D	*A*	13.452 ± 0.205	337	13.730 ± 0.242	93	*0.124*
C18:1 *trans*-11	*A*	none	2.519 ± 0.125	53	2.635 ± 0.099	377	0.277
	B	none	2.619 ± 0.100	302	2.588 ± 0.107	128	0.695
	C	none	2.610 ± 0.098	357	2.599 ± 0.120	73	0.905
	D	none	2.629 ± 0.098	337	2.540 ± 0.115	93	0.302
C18:2 *trans*-9,12	*A*	none	0.345 ± 0.010	53	0.353 ± 0.008	377	0.330
	B	none	0.356 ± 0.008	302	0.344 ± 0.009	128	*0.056*
	C	none	0.348 ± 0.008	357	0.361 ± 0.010	73	*0.090*
	D	none	0.352 ± 0.008	337	0.350 ± 0.009	93	0.791
	B	*C*	0.357 ± 0.009	302	0.345 ± 0.010	128	*0.076*
	C	*B*	0.348 ± 0.009	357	0.359 ± 0.011	73	*0.136*
C18:2 *cis*-9, *trans*-12	*A*	none	0.071 ± 0.003	53	0.070 ± 0.002	377	0.786
	B	none	0.071 ± 0.002	302	0.069 ± 0.003	128	0.373
	C	none	0.070 ± 0.002	357	0.070 ± 0.003	73	0.973
	D	none	0.070 ± 0.002	337	0.072 ± 0.003	93	0.312
C18:2 *trans*-9, *cis*-12	*A*	none	0.438 ± 0.021	53	0.453 ± 0.016	377	0.408
	B	none	0.448 ± 0.016	302	0.450 ± 0.018	128	0.882
	C	none	0.447 ± 0.016	357	0.457 ± 0.020	73	0.540
	D	none	0.454 ± 0.016	337	0.436 ± 0.019	93	0.203
C18:2 *cis*-9, *trans*-13	*A*	none	0.290 ± 0.006	53	0.285 ± 0.005	377	0.306
	B	none	0.289 ± 0.005	302	0.282 ± 0.005	128	*0.062*
	C	none	0.284 ± 0.005	357	0.293 ± 0.006	73	*0.064*
	D	none	0.285 ± 0.005	337	0.291 ± 0.006	93	*0.167*
	B	*C, D*	0.291 ± 0.006	302	0.285 ± 0.006	128	*0.118*
	C	*B, D*	0.285 ± 0.006	357	0.294 ± 0.007	73	*0.068*
	D	*B, C*	0.286 ± 0.006	337	0.292 ± 0.007	93	*0.137*

(Continue on next page)

(Continued)

C18:2 *cis*-9,12	*A*	none	0.707 ± 0.014	53	0.691 ± 0.011	377	*0.199*
	B	none	0.700 ± 0.011	302	0.686 ± 0.012	128	*0.098*
	C	none	0.692 ± 0.011	357	0.703 ± 0.014	73	0.315
	D	none	0.692 ± 0.011	337	0.704 ± 0.013	93	*0.186*
	A	*B, D*	0.710 ± 0.017	53	0.688 ± 0.014	377	*0.076*
	B	*A, D*	0.706 ± 0.016	302	0.688 ± 0.016	128	**0.042**
	D	*A, B*	0.693 ± 0.015	337	0.703 ± 0.016	93	*0.359*
C18:3 *cis*-9,12,15	*A*	none	0.800 ± 0.020	53	0.778 ± 0.015	377	0.203
	B	none	0.792 ± 0.016	302	0.769 ± 0.017	128	*0.061*
	C	none	0.778 ± 0.015	357	0.800 ± 0.019	73	*0.141*
	D	none	0.781 ± 0.015	337	0.791 ± 0.018	93	0.425
	B	*C*	0.793 ± 0.016	302	0.771 ± 0.017	128	*0.070*
	C	*B*	0.777 ± 0.017	357	0.796 ± 0.020	73	0.204
C19:0	*A*	none	1.133 ± 0.005	53	0.135 ± 0.004	377	0.574
	B	none	0.135 ± 0.004	302	0.134 ± 0.004	128	0.629
	C	none	0.135 ± 0.004	357	0.134 ± 0.004	73	0.951
	D	none	0.134 ± 0.004	337	0.135 ± 0.004	93	0.994
C20:0	*A*	none	0.125 ± 0.003	53	0.125 ± 0.002	377	0.968
	B	none	0.125 ± 0.002	302	0.124 ± 0.003	128	0.851
	C	none	0.124 ± 0.002	357	0.127 ± 0.003	73	*0.275*
	D	none	0.125 ± 0.002	337	0.124 ± 0.003	93	0.626
C22:0	*A*	none	0.065 ± 0.002	53	0.063 ± 0.002	377	0.444
	B	none	0.063 ± 0.002	302	0.065 ± 0.002	128	0.242
	C	none	0.063 ± 0.002	357	0.065 ± 0.002	73	0.371
	D	none	0.063 ± 0.002	337	0.065 ± 0.002	93	0.333
C20:1 *cis*-5	*A*	none	0.056 ± 0.003	53	0.057 ± 0.002	377	0.597
	B	none	0.057 ± 0.002	302	0.056 ± 0.002	128	0.739
	C	none	0.056 ± 0.002	357	0.057 ± 0.002	73	0.608
	D	none	0.057 ± 0.002	337	0.056 ± 0.002	93	0.647
C20:1 *cis*-9	*A*	none	0.147 ± 0.004	53	0.152 ± 0.003	377	*0.151*
	B	none	0.151 ± 0.003	302	0.150 ± 0.004	128	0.798
	C	none	0.152 ± 0.003	357	0.147 ± 0.004	73	*0.145*
	D	none	0.150 ± 0.003	337	0.152 ± 0.004	93	0.575
	A	*C*	0.147 ± 0.004	53	0.152 ± 0.004	377	*0.169*
	C	*A*	0.152 ± 0.004	357	0.147 ± 0.004	73	*0.167*
C20:1 *cis*-11	*A*	none	0.077 ± 0.002	53	0.076 ± 0.002	377	0.687
	B	none	0.077 ± 0.002	302	0.076 ± 0.002	128	0.354
	C	none	0.076 ± 0.002	357	0.077 ± 0.002	73	0.823
	D	none	0.077 ± 0.002	337	0.076 ± 0.002	93	0.504
C18:3 *cis*-6,9,12	*A*	none	0.075 ± 0.002	53	0.073 ± 0.002	377	0.331
	B	none	0.074 ± 0.002	302	0.073 ± 0.002	128	0.462
	C	none	0.074 ± 0.002	357	0.075 ± 0.002	73	0.514
	D	none	0.074 ± 0.002	337	0.073 ± 0.002	93	0.635
C20:3 *cis*-8,11,14	*A*	none	0.031 ± 0.001	53	0.030 ± 0.001	377	0.412
	B	none	0.030 ± 0.001	302	0.030 ± 0.001	128	0.242
	C	none	0.029 ± 0.001	357	0.031 ± 0.001	73	**0.017**
	D	none	0.030 ± 0.001	337	0.030 ± 0.001	93	0.478
C22:1 *trans*-13	*A*	none	0.068 ± 0.002	53	0.067 ± 0.002	377	0.706
	B	none	0.067 ± 0.002	302	0.068 ± 0.002	128	0.393
	C	none	0.066 ± 0.002	357	0.071 ± 0.002	73	**0.004**
	D	none	0.067 ± 0.002	337	0.066 ± 0.002	93	0.450

(Continue on next page)

(Continued)

C20:4 *cis*-5,8,11,14	*A*	none	0.035 ± 0.001	53	0.034 ± 0.001	377	0.275
	B	none	0.034 ± 0.001	302	0.034 ± 0.001	128	0.334
	C	none	0.034 ± 0.001	357	0.034 ± 0.001	73	0.749
	D	none	0.034 ± 0.001	337	0.034 ± 0.001	93	0.638
C20:5 *cis*-5,8,11,14,17	*A*	none	0.089 ± 0.002	53	0.086 ± 0.002	377	*0.166*
	B	none	0.087 ± 0.002	302	0.088 ± 0.002	128	0.270
	C	none	0.086 ± 0.002	357	0.089 ± 0.002	73	*0.131*
	D	none	0.087 ± 0.002	337	0.087 ± 0.002	93	0.687
	A	*C*	0.089 ± 0.002	53	0.087 ± 0.002	377	0.206
	C	*A*	0.087 ± 0.002	357	0.089 ± 0.002	73	*0.134*
C22:5 *cis*-7,10,13,16,19	*A*	none	0.125 ± 0.004	53	0.122 ± 0.003	377	0.436
	B	none	0.122 ± 0.003	302	0.123 ± 0.004	128	0.712
	C	none	0.122 ± 0.003	357	0.122 ± 0.004	73	0.853
	D	none	0.122 ± 0.003	337	0.125 ± 0.004	93	0.296
C24:0	*A*	none	0.045 ± 0.002	53	0.044 ± 0.001	377	0.240
	B	none	0.043 ± 0.001	302	0.045 ± 0.001	128	*0.060*
	C	none	0.044 ± 0.001	357	0.045 ± 0.002	73	0.507
	D	none	0.044 ± 0.001	337	0.045 ± 0.001	93	0.331
CLA *cis*-9, *trans*-11	*A*	none	0.936 ± 0.051	53	0.985 ± 0.040	377	0.265
	B	none	0.985 ± 0.041	302	0.953 ± 0.043	128	0.312
	C	none	0.970 ± 0.040	357	0.983 ± 0.049	73	0.747
	D	none	0.978 ± 0.040	337	0.957 ± 0.047	93	0.549

[1]Predicted means and standard error of those means derived from GLMM. Cow age, herd, days in milk (DIM), and myostatin variants were fitted to the models as fixed effects, with DIM and age fitted as covariates. $0.05 < P < 0.2$ in italics and $P < 0.05$ in bold.

Table 3.2 Associations between bovine *MSTN* variants and grouped milk FAs in NZ HF × J-cross cows

Grouped Fatty Acids	Variant	Other variants in model	Mean FAME level ± SE[1] (g/100 g milk FA)				*P*
			Absent	n	Present	n	
SCFA	*A*	None	4.002 ± 0.048	53	4.020 ± 0.037	377	0.531
	B	None	4.031 ± 0.037	302	4.001 ± 0.040	128	0.252
	C	None	4.022 ± 0.036	357	4.011 ± 0.045	73	0.871
	D	None	4.021 ± 0.036	337	4.001 ± 0.043	93	0.400
MCFA	*A*	None	19.642 ± 0.274	53	19.688 ± 0.216	377	0.841
	B	None	19.732 ± 0.218	302	19.589 ± 0.233	128	0.397
	C	None	19.633 ± 0.215	357	19.811 ± 0.261	73	0.397
	D	None	19.783 ± 0.212	337	19.352 ± 0.251	93	0.021
LCFA	*A*	None	48.263 ± 0.458	53	48.281 ± 0.361	377	0.962
	B	None	48.133 ± 0.364	302	48.510 ± 0.389	128	*0.182*
	C	None	48.386 ± 0.359	357	47.944 ± 0.436	73	0.207
	D	None	48.198 ± 0.357	337	48.518 ± 0.421	93	0.308
Total C18:2	*A*	None	2.787 ± 0.077	53	2.837 ± 0.061	377	0.446
	B	None	2.850 ± 0.061	302	2.785 ± 0.066	128	*0.171*
	C	None	2.811 ± 0.060	357	2.866 ± 0.074	73	0.349
	D	None	2.830 ± 0.060	337	2.810 ± 0.071	93	0.703
Total C18:3	*A*	None	0.875 ± 0.020	53	0.852 ± 0.016	377	*0.182*
	B	None	0.866 ± 0.016	302	0.843 ± 0.017	128	*0.057*
	C	None	0.851 ± 0.016	357	0.875 ± 0.019	73	*0.132*
	D	None	0.855 ± 0.016	337	0.865 ± 0.019	93	0.466
	A	*B, C*	0.880 ± 0.025	53	0.846 ± 0.022	377	*0.053*
	B	*A, C*	0.876 ± 0.023	302	0.846 ± 0.023	128	**0.020**
	C	*A, B*	0.856 ± 0.023	357	0.871 ± 0.025	73	0.326
Omega 3	*A*	None	1.013 ± 0.020	53	0.986 ± 0.016	377	*0.125*
	B	None	1.000 ± 0.016	302	0.980 ± 0.017	128	*0.107*
	C	None	0.987 ± 0.016	357	1.011 ± 0.019	73	*0.124*
	D	None	0.990 ± 0.016	337	1.003 ± 0.019	93	0.345
	A	*B, C*	1.018 ± 0.024	53	0.981 ± 0.021	377	**0.041**
	B	*A, C*	1.011 ± 0.024	302	0.984 ± 0.025	128	**0.036**
	C	*A, B*	0.992 ± 0.023	357	1.009 ± 0.025	73	0.292
Omega 6	*A*	None	0.847 ± 0.015	53	0.828 ± 0.012	377	*0.141*
	B	None	0.838 ± 0.012	302	0.824 ± 0.013	128	*0.140*
	C	None	0.829 ± 0.012	357	0.843 ± 0.015	73	0.245
	D	None	0.829 ± 0.012	337	0.841 ± 0.014	93	0.247
	A	*B*	0.850 ± 0.018	53	0.824 ± 0.015	377	*0.056*
	B	*A*	0.845 ± 0.018	302	0.826 ± 0.018	128	*0.055*

(Continue on next page)

(Continued)

PUFA	*A*	None	3.940 ± 0.082	53	3.960 ± 0.065	377	0.776
	B	None	3.987 ± 0.065	302	3.903 ± 0.070	128	*0.094*
	C	None	3.935 ± 0.064	357	4.017 ± 0.078	73	0.193
	D	None	3.957 ± 0.064	337	3.949 ± 0.076	93	0.891
	B	*C*	3.988 ± 0.066	302	3.903 ± 0.070	128	*0.095*
	C	*B*	3.932 ± 0.070	357	4.005 ± 0.084	73	0.251
branched FA	*A*	None	1.580 ± 0.026	53	1.593 ± 0.020	377	0.558
	B	None	1.598 ± 0.021	302	1.576 ± 0.022	128	0.155
	C	None	1.583 ± 0.020	357	1.608 ± 0.025	73	0.209
	D	None	1.595 ± 0.020	337	1.572 ± 0.024	93	0.196
	B	*D*	1.595 ± 0.023	302	1.571 ± 0.025	128	0.126
	D	*B*	1.594 ± 0.023	337	1.569 ± 0.026	93	0.156
C12:1 index	*A*	None	2.231 ± 0.062	53	2.289 ± 0.049	377	0.276
	B	None	2.299 ± 0.049	302	2.237 ± 0.053	128	0.104
	C	None	2.280 ± 0.049	357	2.261 ± 0.059	73	0.686
	D	None	2.281 ± 0.048	337	2.257 ± 0.057	93	0.573
C16:1 index	*A*	None	3.516 ± 0.097	53	3.498 ± 0.077	377	0.831
	B	None	3.529 ± 0.078	302	3.459 ± 0.083	128	0.243
	C	None	3.495 ± 0.076	357	3.523 ± 0.093	73	0.709
	D	None	3.474 ± 0.076	337	3.589 ± 0.089	93	*0.083*

[1]Predicted means and standard error of those means derived from GLMM. ‘Cow age’, ‘Herd’, ‘days in milk (DIM)’, and ‘myostatin variants were fitted to the models as fixed effects, with DIM and age fitted as covariates. $0.05 < P < 0.20$ in italics and $P < 0.05$ in bold.

Table 3.3 Associations between *MSTN* variants and gross milk production traits in NZ HF × J-cross cows

			Mean ± SE[1]				
	Variant	**Other Variants in model**	**Absent**	**n**	**Present**	**n**	***P***
Average milk Volume(L)	*A*	none	23.672 ± 0.580	53	22.961 ± 0.457	377	0.151
	B	none	22.888 ± 0.462	302	23.521 ± 0.493	128	0.077
	C	none	23.128 ± 0.456	357	23.137 ± 0.555	73	0.984
	D	none	23.153 ± 0.454	337	23.060 ± 0.535	93	0.816
	A	*B*	23.619 ± 0.610	53	23.017 ± 0.495	377	0.235
	B	*A*	22.889 ± 0.462	302	23.522 ± 0.493	128	0.077
Fat (%)	*A*	none	4.834 ± 0.090	53	4.975 ± 0.071	377	0.064
	B	none	4.944 ± 0.072	302	4.937 ± 0.077	128	0.892
	C	none	4.973 ± 0.070	357	4.847 ± 0.086	73	0.068
	D	none	4.955 ± 0.070	337	4.900 ± 0.083	93	0.368
	A	*C*	4.829 ± 0.098	53	4.958 ± 0.082	377	0.094
	C	*A*	4.952 ± 0.085	357	4.839 ± 0.097	73	0.102
Protein (%)	*A*	none	3.946 ± 0.048	53	4.017 ± 0.037	377	0.079
	B	none	4.016 ± 0.038	302	3.974 ± 0.041	128	0.154
	C	none	4.015 ± 0.037	357	3.952 ± 0.045	73	0.084
	D	none	4.009 ± 0.037	337	3.971 ± 0.044	93	0.244
	A	*B, C*	3.944 ± 0.053	53	4.005 ± 0.045	377	0.145
	B	*A, C*	4.004 ± 0.051	302	3.954 ± 0.053	128	0.088
	C	*A, B*	4.011 ± 0.044	357	3.941 ± 0.051	73	0.058

[1]Predicted means and standard error of those means were derived from the GLMMs. Cow age, herd, and days in milk (DIM) were fitted as fixed effects, with DIM and age fitted as covariates.

Table 3.4 Associations between milk FA levels and *MSTN* genotypes in HF × J-cross cows

Individual/Grouped FAs	Mean FAME level ± SE[1] (g/100 g milk FA)				*p*
	AA	***AB***	***AC***	***AD***	
	n = 151	n = 92	n = 53	n = 65	
C4:0	1.252 ± 0.023	1.243 ± 0.024	1.236 ± 0.026	1.258 ± 0.026	0.776
C6:0	1.551 ± 0.020	1.536 ± 0.020	1.531 ± 0.022	1.532 ± 0.022	0.489
C8:0	1.195 ± 0.017	1.174 ± 0.017	1.184 ± 0.018	1.164 ± 0.019	0.104
C10:0	3.293 ± 0.066[a]	3.226 ± 0.067[ab]	3.290 ± 0.072[ab]	3.144 ± 0.073[b]	0.040
C10:1	0.290 ± 0.008	0.277 ± 0.008	0.282 ± 0.009	0.279 ± 0.009	*0.091*
C11:0	0.060 ± 0.003	0.056 ± 0.003	0.055 ± 0.004	0.055 ± 0.004	0.105
C12:0	4.014 ± 0.088[a]	3.906 ± 0.089[ab]	4.011 ± 0.097[ab]	3.788 ± 0.098[b]	0.012
C13:0 *iso*	0.083 ± 0.003	0.078 ± 0.003	0.080 ± 0.003	0.077 ± 0.003	0.043
C12:1	0.095 ± 0.003[a]	0.089 ± 0.003[ab]	0.093 ± 0.004[ab]	0.087 ± 0.004[b]	0.005
C13:0 *anteiso*	0.038 ± 0.001	0.037 ± 0.001	0.038 ± 0.001	0.038 ± 0.001	0.323
C13:0	0.121 ± 0.005	0.116 ± 0.005	0.115 ± 0.005	0.115 ± 0.005	0.291
C14:0	12.494 ± 0.151	12.440 ± 0.154	12.670 ± 0.167	12.324 ± 0.169	0.176
C14:1 *cis*-9	0.995 ± 0.039	0.947 ± 0.040	0.974 ± 0.043	0.954 ± 0.044	0.331
C15:0 *iso*	0.297 ± 0.005	0.290 ± 0.005	0.299 ± 0.005	0.289 ± 0.005	0.051
C15:0 *anteiso*	0.632 ± 0.017	0.625 ± 0.017	0.638 ± 0.019	0.635 ± 0.019	0.859
C15:0	1.462 ± 0.030	1.448 ± 0.031	1.454 ± 0.034	1.441 ± 0.034	0.838
C15:1	0.281 ± 0.006	0.279 ± 0.006	0.284 ± 0.006	0.283 ± 0.006	0.767
C16:0	36.757 ± 0.541	37.021 ± 0.551	36.440 ± 0.598	36.952 ± 0.606	0.707
C16:1 *cis*-9	1.327 ± 0.044	1.295 ± 0.044	1.345 ± 0.048	1.369 ± 0.049	0.295
C17:0 *iso*	0.556 ± 0.012	0.552 ± 0.012	0.561 ± 0.013	0.552 ± 0.013	0.855
C17:0	0.695 ± 0.012	0.705 ± 0.012	0.696 ± 0.013	0.695 ± 0.013	0.678
C17:1	0.209 ± 0.004	0.207 ± 0.005	0.213 ± 0.005	0.213 ± 0.005	0.404
C18:0	8.565 ± 0.226	8.708 ± 0.230	8.567 ± 0.250	8.626 ± 0.253	0.844
C18:1 *trans*-11	2.705 ± 0.130	2.732 ± 0.132	2.749 ± 0.144	2.748 ± 0.146	0.971
C18:1 *cis*-9	13.540 ± 0.267	13.599 ± 0.272	13.593 ± 0.295	13.883 ± 0.299	0.480
C18:2 *trans*-9,12	0.358 ± 0.011	0.352 ± 0.011	0.368 ± 0.012	0.364 ± 0.012	0.429
C18:2 *cis*-9, *trans*-13	0.294 ± 0.006	0.285 ± 0.006	0.297 ± 0.007	0.297 ± 0.007	0.104
C18:2 *cis*-9, *trans*-12	0.071 ± 0.003	0.068 ± 0.003	0.070 ± 0.004	0.071 ± 0.004	0.590
C18:2 *trans*-9, *cis*-12	0.476 ± 0.021	0.484 ± 0.022	0.495 ± 0.024	0.474 ± 0.024	0.738
C18:2 *cis*-9,12	0.696 ± 0.014	0.679 ± 0.014	0.698 ± 0.016	0.701 ± 0.016	0.266
C19:0	0.135 ± 0.005	0.135 ± 0.005	0.137 ± 0.005	0.137 ± 0.005	0.958
C18:3 *cis*-6,9,12	0.075 ± 0.002	0.073 ± 0.002	0.074 ± 0.002	0.073 ± 0.002	0.513
C18:3 *cis*-9,12,15	0.793 ± 0.020	0.763 ± 0.021	0.802 ± 0.022	0.787 ± 0.023	0.142
CLA *cis*-9, *trans*-11	1.032 ± 0.053	1.016 ± 0.054	1.057 ± 0.058	1.052 ± 0.059	0.832
C20:0	0.123 ± 0.003	0.124 ± 0.003	0.126 ± 0.003	0.124 ± 0.004	0.834
C20:1 *cis*-5	0.059 ± 0.003	0.060 ± 0.003	0.061 ± 0.003	0.060 ± 0.003	0.882
C20:1 *cis*-9	0.153 ± 0.005	0.154 ± 0.005	0.151 ± 0.005	0.153 ± 0.005	0.926
C20:1 *cis*-11	0.078 ± 0.002	0.075 ± 0.002	0.077 ± 0.003	0.073 ± 0.003	0.120
C20:3 *cis*-8,11,14	0.029 ± 0.001	0.029 ± 0.001	0.030 ± 0.001	0.028 ± 0.001	0.109
C20:4 *cis*-5,8,11,14	0.033 ± 0.001	0.033 ± 0.001	0.033 ± 0.001	0.032 ± 0.001	0.957
C22:0	0.062 ± 0.002	0.064 ± 0.002	0.063 ± 0.003	0.063 ± 0.003	0.465
C22:1 *trans*-13	0.065 ± 0.003	0.068 ± 0.003	0.069 ± 0.003	0.066 ± 0.003	0.169
C20:5 *cis*-5,8,11,14,17	0.085 ± 0.002	0.086 ± 0.002	0.086 ± 0.002	0.085 ± 0.002	0.914
C24:0	0.043 ± 0.002	0.046 ± 0.002	0.043 ± 0.002	0.044 ± 0.002	0.214
C22:5 *cis*-7,10,13,16,19	0.118 ± 0.004	0.117 ± 0.004	0.118 ± 0.005	0.118 ± 0.005	0.994
SCFA	3.998 ± 0.051	3.952 ± 0.052	3.950 ± 0.056	3.953 ± 0.057	0.505
MCFA	19.801 ± 0.279	19.572 ± 0.284	19.971 ± 0.309	19.256 ± 0.313	*0.052*
LCFA	47.842 ± 0.475	48.250 ± 0.484	47.525 ± 0.525	48.084 ± 0.532	0.410
Total C18:1	16.987 ± 0.315	17.071 ± 0.321	17.083 ± 0.348	17.374 ± 0.353	0.531

(Continue on next page)

(Continued)

Total C18:2	2.927 ± 0.079	2.884 ± 0.080	2.985 ± 0.087	2.959 ± 0.088	0.561
Total C18:3	0.868 ± 0.021	0.836 ± 0.021	0.876 ± 0.023	0.860 ± 0.023	0.122
Omega 3	0.996 ± 0.021	0.966 ± 0.021	1.007 ± 0.023	0.990 ± 0.024	0.159
Omega 6	0.832 ± 0.015	0.814 ± 0.015	0.835 ± 0.017	0.834 ± 0.017	0.308
MUFA	20.538 ± 0.329	20.521 ± 0.335	20.633 ± 0.363	20.909 ± 0.368	0.541
PUFA	4.059 ± 0.084	3.985 ± 0.085	4.128 ± 0.092	4.082 ± 0.094	0.319
branched FA	1.605 ± 0.027	1.582 ± 0.027	1.616 ± 0.030	1.591 ± 0.030	0.508
Total UFA	24.597 ± 0.391	24.507 ± 0.397	24.762 ± 0.432	24.991 ± 0.437	0.538
Total SFA	71.822 ± 0.427	71.946 ± 0.434	71.617 ± 0.471	71.463 ± 0.478	0.604
unsaturated index	25.517 ± 0.414	25.413 ± 0.421	25.697 ± 0.457	25.915 ± 0.463	0.558
C10:1 index	8.190 ± 0.259	7.967 ± 0.101	7.991 ± 0.110	8.256 ± 0.290	0.496
C12:1 index	2.319 ± 0.065	2.219 ± 0.066	2.281 ± 0.072	2.241 ± 0.073	0.170
C14:1 index	7.375 ± 0.285	7.062 ± 0.290	7.146 ± 0.315	7.185 ± 0.319	0.473
C16:1 index	3.481 ± 0.099	3.374 ± 0.101	3.542 ± 0.110	3.551 ± 0.111	0.172
C18:1 index	66.519 ± 0.634	66.234 ± 0.646	66.691 ± 0.701	66.885 ± 0.710	0.706
CLA index	27.075 ± 0.518	26.668 ± 0.528	27.345 ± 0.573	27.355 ± 0.581	0.415

[1]Predicted means and standard error of those means derived from GLMM. ‘Cow age’, ‘Herd’, ‘days in milk (DIM)’, and ‘myostatin genotypes’ were fitted to the models as fixed effects, with DIM and age fitted as covariates. $P<0.05$ in bold. Means within a row that do not share a superscript letter are separated by Bonferroni test at $P<0.05$

3.3 Discussion

All seven of the nucleotide substitutions identified in this study have been previously reported in a study of New Zealand cattle breeds, which included; Hereford, Angus, Shorthorn, Charolais, Red Poll, South Devon, Simmental, Murray Grey, NZ HF × J–cross cattle and some composite breeds (Haruna *et al.*, 2020). The nucleotide substitutions c.373+751G/T, c.373+803T/G, c.373+877A/G, c.373+895G/C and c.374-909C/T were identified in all the ten aforementioned breeds, while c.374-842G/C was found in all but four breeds (Red Poll, Shorthorn, Simmental and Composites breeds) and the c.374-812A/G was only found in Shorthorn and NZ HF × J-cross cows.

This is the first study demonstrating association of *MSTN* variants with milk production traits and composition of milk FAs in NZ HF × J-cross cattle. Overall, the results for associations in the single variant model revealed that the presence of variant *D* was associated with lower adjusted mean levels of some SFAs, whereas *C* was associated with higher adjusted mean levels of two UFAs and *B* was associated with lower level of saturated branched FA.

In a previous study, Smith *et al.* (2000) identified a mutant '*mh*' allele of bovine *MSTN* with a 11-bp deletion in the exon 3 region. This was described in British South Devon cattle, and it was the same allele first identified in the Belgian Blue breed, causing the 'double-muscled' phenotype. Also, in a later investigation of *MSTN* in 146 British South Devon cattle, Wiener *et al.* (2009), showed that this same mutant '*mh*' allele, reduced the levels of total SFA and total MUFA in the muscle ($P < 0.05$). Their report also revealed an increase in the ratio of PUFA: SFA in total lipid to be greater in *mh/mh* individuals than in the other two genotype (*mh/+* and *+/+)* classes ($P < 0.001$). This suggested that the *'mh*' allele is associated with reduced fat levels, particularly with the levels of SFA and MUFA. The increase in the ratio of PUFA to SFA would be expected, especially when the concentration of muscle fat declines.

In a similar report on associations between the '*mh*' allele carrying the 11-bp deletion with intramuscular FA composition in *MSTN* of Belgian Blue young bulls, Raes *et al.* (2001) revealed that animals with the +/ + (normal) genotype showed a higher relative amount of the SFA; C14:0 and C16:0 and a higher relative amount of all MUFA C16:1; C17:1 and C18:1, but, the relative proportion of PUFA in total FAs significantly increased with increasing mutant '*mh*' alleles. In the current study, the genotype model revealed that cows carrying the *AA* (normal/predominant) genotype showed an increase in the amount of two medium chain-length SFAs (C10:0 and C12:0) and one MUFA (C12:1), but those carrying *MSTN* variant heterozygous genotype *AD* showed decreased levels of the same FAs above. Cows with the *AC* genotype showed a relative increase in the level of the PUFA (C18:3 *cis*-9, 12, 15) from 0.793 ± 0.020 to 0.802 ± 0.022, even though no association was revealed for the *AC* genotype ($P = 0.142$). Although the current study did not identify the '*mh*' allele reported earlier, it revealed that variants *B* and *D* in the intronic region of *MSTN* decreased some of the SFAs content in milk. While it may be difficult to conclude that the findings in the current study are similar to the work

of Weiner *et al.* (2009) and Raes *et al.* (2001), especially because it investigated the intronic region of *MSTN* in non-doubled muscled cattle breed, it is important to note that variations in the intronic regions are equally capable of influencing gene expression and/or altering the functionality of a gene (He *et al.*, 2010). In a previous investigation, He *et al.* (2010) transformed C2C12 cell-lines with a transgene construct that contained bovine *MSTN* promoter (pMD-MSTNPro) and another construct containing the first intron of bovine *MSTN* (pMD-Intron1). They observed an increase in the mean fluorescence intensity of Green Fluorescent Protein (GFP) gene and percentage of fluorescence positive cells and concluded that the presence of intron 1 of bovine *MSTN* improved the expression of GFP in the transformed cells.

Although the effect of variant *A* (the most common variant), and variant *C* on gross milk production traits were not significant (Table 3.3), an interesting trend is still evident as *A* was showing a trend of association with higher total percentage of fat ($P = 0.064$) and protein ($P = 0.079$), while *C* decreased total percentage of fat ($P = 0.068$) and protein ($P = 0.084$). These results, combined with the aforementioned effects of *C* in increasing the PUFA (C20:3 *cis*-8, 11, 14) ($P = 0.017$) and *D* in decreasing grouped MCFA ($P = 0.021$) are in agreement with the findings of (Buske *et al.*, 2011). Buske *et al.* (2011) suggested that one copy of the so-called 'wild-type + allele' of *MSTN* is responsible for moderately higher milk, protein, and fat yields in dual purpose Belgian Blue (DP-BB) cows, whereas a single copy of the mutant '*mh*' allele decreases the SFA contents in milk for DP-BB. This indicates that, there is a possible superior effect of the '+' allele compared with the '*mh*' allele in terms of increasing milk, fat, and protein yields. Buske *et al.* (2011) further suggested that, selection of the '+' allele has the potential to increase conventional milk production traits in the dual-purpose Belgian Blue breed. The current investigation appears to have substantiated the claim that the '+' allele could be selected for increased percentage of fat and protein, whereas, if the focus is on milk FA composition, the '*mh*' allele should be selected for reduced SFA levels and increased PUFA in milk. Even though the current study did not involve the *MSTN* '*mh*' allele responsible for the double-muscling trait in DP-BB breed, it has provided more enlightenment on the possible antagonistic effect of the wild type *MSTN* allele in improving milk fat and protein, and the variant *MSTN* allele in reducing SFA levels and increasing PUFA levels in milk.

Since the observed *MSTN* variations in the current study occurred in a non-expressed region of the gene, it is perhaps less obvious how they might affect the structure or function of the protein. They might however influence the rate of transcription and/or translation of the gene, as was described by Liu *et al.* (1995) in an investigation that involved both transgenic mice and *in-vitro* studies. In that study, intron sequences in the transgene that encoded rat growth hormone, were observed to stimulate transcription by promoting assembly of an ordered nucleosome array in the vicinity of the promoter. Additionally, while the observed intron variation does not yield any amino acid change, several studies have shown that synonymous nucleotide substitutions can affect the phenotypic characteristics of the protein product by altering mRNA structure, protein stability, the electrical charge of the resulting

polypeptide, and codon usage during mRNA translation (Sauna and Kimchi-Sarfaty, 2011; Hunt *et al.*, 2014).

Previous reports on the differences in the composition of milk FAs between Holstein-Friesian and the Jersey breeds, suggest that milk from Jersey cows tends to have higher concentrations of some short- and medium-chain SFAs, but lower concentrations of some UFAs (Arnould and Soyeurt, 2009). It is therefore interesting to note that the current investigation of NZ HF × J-cross cows, revealed that variant *D* was associated with a decrease in the amount of SFA in milk, whereas variant *C* was associated with increased UFA content. This discovery could be of benefit in terms of its potential application in cross-breeding and gene marker development, particularly in selection for decreased SFA levels and increased UFA levels in milk.

3.4 Conclusion

The current findings suggest that *MSTN* variant *D* is associated with a decrease in SFA levels, while variant *C* is associated with an increase in UFA content in the milk of NZ HF × J-cross cows. Cows with the *CD* genotype might therefore produce a 'preferred' FA profile in milk. However, because there were insufficient cattle with the homozygous genotypes *BB, CC* and *DD*, or the heterozygous genotypes (*BC*, *BD* and *CD*) in the samples investigated, further investigation involving larger sample sizes across different farms and breeds is needed to validate this claim.

Chapter 4

Identification of genetic variations in bovine leptin gene (*LEP*) and its association with milk production traits, bodyweight and composition of milk fatty acids in NZ HF × J-cross cows

One of the most important factors affecting animal productivity is feed intake. Poor feed intake or inadequate nutrition can affect bodyweight, growth, reproduction, and milk production, and it can also decrease immunity. Accordingly, a desire to increase animal productivity may require the adoption of measures aimed at increasing or maximizing feed intake. To achieve this, understanding the role of genetics in regulating this important trait, is important.

Recently, there has been a growing interest in genomic selection programmes aimed at modifying milk traits and the composition of milk FAs using the candidate gene approach. Several genes have been identified to affect milk traits and FA composition, and one of such genes is the leptin (*LEP*) gene.

Leptin is a product of the gene that is now called *LEP*, but previously was known as *OB*, *OBS* and *LEPD*. Bovine *LEP* has been mapped to chromosome 4 (Pomp *et al.*, 1997) and it consists of three exons separated by two introns. Exon 1 and part of exon 2 (four nucleotides) are not translated, and only the remaining part of exons 2 and 3 are translated into the functional 16-kDa leptin protein of 146 amino acids in length.

Leptin is considered to be a protein hormone and is mainly secreted from white adipose tissue. This protein is found to regulate feed intake, bodyweight, immune function, and reproduction (Santos-Alvarez *et al.*, 1999; Kadokawa *et al.*, 2000; Block *et al.*, 2001). Leptin acts on receptors in the lateral hypothalamus to inhibit hunger (Elias *et al.*, 1999) and in the medial hypothalamus to stimulate satiety (Elmquist *et al.*, 1999). It counteracts the effects of neuropeptide Y, a potent hunger promoter secreted by cells in the gut and in the hypothalamus. The hypothalamic release of neuropeptide Y eventually results in a decrease in feed intake and an increase in energy expenditure, among other things (Houseknecht *et al.*, 1998).

Several nucleotide sequence variations have been reported in bovine *LEP*, with these including microsatellite repeat number variation and nucleotide sequence variation (Stone *et al.*, 1996b; Konfortov *et al.*, 1999). Konfortov *et al.* (1999), investigated a 1,788 bp portion of *LEP* (comprising of exons 2 and 3, and parts of introns 1 and 2) in 13 different breeds of *Bos taurus* and *Bos indicus* cattle. They reported 20 nucleotide variations, six of which occurred in the exons and at a frequency of approximately one per 84 bp. The frequency of the intron sequence variations was one per 92 bp, giving an overall frequency of variation of one per 89 bp. This suggests the gene is polymorphic, and that if more cattle breeds are examined, more variation might be detected.

Previous reports with association analyses have highlighted the role of leptin gene polymorphisms on the concentration of circulating leptin (Liefers *et al.*, 2005; Jonas *et al.*, 2016), as well as its effects on some traits of economic value such as the yield and quality of meat and milk obtained from farm animal species. For example, a significant effect of leptin gene polymorphism was found on weaning weight in sheep (Hajihosseinlo *et al.*, 2012). In cattle, leptin or leptin receptor gene polymorphisms have been associated with carcass FA composition (Kawaguchi *et al.*, 2017), milk yield (Banos *et al.*, 2008), milk fat (Giblin *et al.*, 2010; De Matteis *et al.*, 2012), milk protein (Giblin *et al.*, 2010), and milk FA composition (Pegolo *et al.*, 2016). However, despite several investigations of leptin gene polymorphisms and their effects on composition of milk FAs and milk production traits in cattle, there is no such report in NZ HF x J-cross cows.

This chapter is aimed at investigating an extended region of bovine *LEP* in a variety of cattle breeds from New Zealand and Nigeria, with a view to identifying unique and novel variations specific to a breed or shared across breeds. This would enable improved understanding of the variability of this gene, and subsequently, may allow identification of potential gene-markers associated with feed intake in cattle. The effect of leptin gene variations on bodyweight, milk production traits, and the composition of milk FAs in NZ HF × J-cross cows will also be described.

4.1 Materials and Methods

4.1.1 Cattle investigated

This study was approved by the Lincoln University Animal Ethics Committee under the provisions of the Animal Welfare Act 1999 (New Zealand Government), irrespective of whether blood was collected in New Zealand or Nigeria.

A total of 657 cattle were investigated, with these being of a variety of breeds found in New Zealand and Nigeria. These included NZ Hereford (n = 23), Angus (n = 23), Shorthorn (n = 18), and NZ HF × J-cross cows (n = 485); and the Nigerian Sokoto Gudali (SG, n = 18), Red Bororo (RB, n = 34), White Fulani (WF, n = 32) and cross-bred Holstein Friesian × White Fulani (HF × WF-cross) cattle (n = 24). Of these eight breeds, the NZ HF × J-cross cows and the HF × WF-cross are bred for milk production, whereas the Hereford, Angus and Shorthorn breeds are farmed primarily for meat production. The Nigerian SG, RB and WF are dual purpose breeds. Unlike the New Zealand breeds which are *Bos taurus* cattle, the Nigerian WF, RB and SG are of Zebu (*Bos indicus*) origin, and are characterized by having a fatty thoracic hump on their shoulders and large dewlap, and by being adapted to dry environmental conditions (Mattioli *et al.*, 2000).

The cows used for association studies were the NZ HF × J-cross cows. These cows are of variable and unknown breed proportion, and of 3 to 13 years of age. The cows were from two herds, all of which

were grazed on pasture (a mixture of perennial ryegrass and white clover) on the Lincoln University Dairy Farm (LUDF; Canterbury, New Zealand). All the cows calved over the months of August-September and were milked twice a day. The weight (kg) for each cow was measured on a weekly basis over a period of 39 weeks.

4.1.2 Milk sample collection

Milk samples for gross milk trait analysis were collected once a month from September to February and the daily milk yield in litres was recorded using Tru-test milk meters (Tru-test Ltd, Auckland, New Zealand) and Fourier-Transform Infra-Red Spectroscopy (MilkoScan FT 120 Foss, Hillerød, Denmark) was used to analyse milk samples for fat percentage (%) and protein percentage (%). Average daily milk yield, and average protein and fat percentages were calculated over the 6 months of milk collection. The milk samples for FA analysis were collected at 148 ± 19 days in milk from each cow in a single afternoon milking in mid-January (the middle of summer in the Southern Hemisphere). These were frozen at -20 ^{o}C, and then freeze-dried, prior to being individually ground to a fine powder for component analysis.

4.1.3 Gas chromatography of the FAs in the milk samples

The FAs were methylated and extracted in n-heptane, before being analysed by Gas Chromatography (GC) as FA methyl esters (FAMEs). The methylation reactions were performed in 10-mL Kimax tubes. Individual freeze-dried and powdered milk samples (0.17 g), were dissolved in 900 µL of n-heptane (100%, AR grade), before 100 µL of internal standard (5 mg/ml of C21:0 methyl ester in n-heptane) and 4.0 mL of 0.5 M NaOH (in 100% anhydrous methanol) were added.

The tubes were vortexed then incubated in a block heater (Ratek Instruments, Australia) at 50°C for 15 minutes. After cooling to room temperature, another 2.0 mL of n-heptane and 2.0 mL of deionized water was added to each tube. After vortexing, the tubes were centrifuged (Megafuge 1.0R, Heraeus, Germany) for 5 minutes at 1500x g. The top layer of n-heptane was transferred into a second Kimax tube and 2.0 mL of n-heptane was added to each of the original tubes. The extraction was repeated and the n-heptane aspirates were then pooled. Anhydrous sodium sulphate (10 mg) was added to the n-heptane extracts, to remove any residual water.

The GC analysis was carried out using a Shimadzu GC-2010 Gas Chromatograph (Shimadzu Corporation, Kyoto, Japan) equipped with a flame ionization detector and an AOC-20i auto sampler. The output was analysed with GC Solution Software (Shimadzu). For analysis, 1 µL of the n-heptane sample extract was injected into a 100 metre GC column (250 µm × 0.25 µm capillary column, CP-Select, Varian) with a 1:60 split ratio. The separation was undertaken with a helium carrier gas and was run for 92 minutes. The temperature of both the injector and detector were set at 250 °C and the thermal profile of the column consisted of 45 °C for 4 minutes, followed by 27 minutes at 175 °C (ramped at 13 °C/minute), 35 minutes at 215 °C (ramped at 4 °C/minute), and a final ‘bake-off’ at 250

°C for 5 minutes (ramped at 25 °C/minute.). The individual FAMEs were identified by the peak retention time compared to commercially obtained external standards (ME61, ME93, BR3, BR2, ME100, GLC411 and GLC463; Laroden AB, Sweden). Quantification of the individual FAMEs was based on peak area assessment and comparison with the internal and external standards. The threshold for peak area determination on the chromatogram was a 500-unit count, with peaks that were under 500-unit count, being ignored. The calculated minimum component of an individual FAME was therefore 0.01 grams per 100 grams of total FA.

After their individual measurement, the FAs were arranged into various groups and indices. These groups were, short chain-length FAs (SCFA) = C4:0 + C6:0 + C8:0; medium chain-length FAs (MCFA) = C10:0 + C12:0 + C14:0; long chain-length FAs (LCFA) = C15:0 + C16:0 + C17:0 + C18:0 + C19:0 + C20:0 + C22:0 + C24:0; omega 3 FAs = C18:3 *cis*-9, 12, 15 + C20:5 *cis*-5,8, 11, 14, 17 + C22:5 *cis*-7, 10, 13, 16, 19; omega 6 FAs = C18:2 *cis*-9, 12 + C18:3 *cis*-6, 9, 12 + C20:3 *cis*-8, 11, 14 + C20:4 cis-5, 8, 11, 14; monounsaturated FAs (MUFA) = C10:1 + C12:1 + C14:1 *cis*-9 + C15:1 + C16:1 *cis*-9 + C17:1 + C18:1 *trans*-11 + C18:1 *cis*-9 + C18:1 *cis*-(10 to 15) + C20:1 *cis*-5 + C20:1 *cis*-9 + C20:1 *cis*-11 + C22:1 *trans*-13; polyunsaturated FAs (PUFA) = C18:2 *trans*-9, 12 + C18:2 *cis*-9,*trans*-13 + C18:2 *cis*-9,*trans*-12 + C18:2 *trans*-9,*cis*-12 + C18:2 *cis*-9, 12 + C18:3 *cis*-6, 9, 12 + C18:3 *cis*-9, 12, 15 + CLA + C20:3 *cis*-8, 11, 14 + C20:4 *cis*-5, 8, 11, 14 + C20:5 *cis*-5, 8, 11, 14, 17 + C22:5 *cis*-7, 10, 13, 16, 19; and total branched FA= C13:0 *iso* + C13:0 *anteiso* + C15:0 *iso* + C15:0 *anteiso* + C17:0 *iso*.

Unsaturated FA indices were also calculated as follows: C12:1 index (C12:1 divided by the sum of C12:0 and C12:1); C16:1 index (C16:1 *cis*-9 divided by the sum of C16:0 and C16:1 *cis*-9) and C18:1 index (C18:1 *cis*-9 divided by the sum of C18:0 and C18:1 *cis*-9). The method is as described by Li *et al.* (2019), with the un-adjusted mean levels in the 402 cows being calculated and used subsequently in the statistical analyses.

4.1.4 Blood sample collection

Using the conventional method of piercing the animal's ear, blood samples were collected from each cattle onto FTA cards (Whatman, Middlesex, UK). The samples were air dried and DNA purification was carried out using a two-step procedure as described by Zhou *et al.* (2006).

4.1.5 PCR primers to amplify the bovine leptin gene (LEP)

Three pairs of primers were designed using DNAMAN 5.0 software. These were designed to amplify three regions of the bovine leptin gene (GenBank accession number U50365.1). The primer pair 1 (5'-gtctttgaggagatgatagcc-3' and 5'-gctgtctttatgccagggg-3') amplified a 443-bp fragment, consisting of exon 2, and parts of introns 1 and 2; the primer pair 2 (5'-agctagtcaggttccacaag-3'and 5'-ggttctgcaagggtattcag-3') amplified a 400-bp fragment of intron 2; and the primer pair 3 (5'-

ttgctctccccttcctcctg-3'and 5'-ctcaggtttcttccctggac-3') amplified a 430-bp fragment, which consisted of exon 3 and part of intron 2.

4.1.6 Polymerase Chain Reaction (PCR)

The PCR reactions for all three regions were performed in 15-µL reactions containing the genomic DNA on a 1.2-mm diameter disc of FTA paper, 0.25 uM for each primer, 150 uM for each dNTP (Eppendorf, Hamburg, Germany), 3.0 mM Mg^{2+}, 0.5 U of *Taq* DNA polymerase (Qiagen, Hilden, Germany), and 1× the reaction buffer supplied with the enzyme. For each of the three regions, PCR amplification was undertaken in Bio-Rad S1000 thermal cyclers (Bio-Rad, Hercules, CA, USA). The thermal cycling conditions included an initial denaturation at 94 ºC for 2 minutes, followed by 35 cycles of 94 ºC for 30 seconds, annealing for 30 seconds at 60 ºC, extension at 72 ºC for 30 seconds, and a final extension step at 72 ºC for 5 minutes.

4.1.7 Single Strand Conformation Polymorphism (SSCP) analysis

The SSCP technique was used to detect genetic variation in the amplicons obtained from the PCR reactions. A 0.7-µL aliquot of the amplicons was added to 7-µL of loading dye containing 10 mM EDTA, 0.025 % bromophenol blue, 0.025 % xylene-cyanol, and 98 % formamide. The samples were then placed on a hot plate already set at 95 ºC for 5 minutes for denaturation, followed by snap-chilling on wet ice. Samples were then loaded onto 16 cm × 18 cm, acrylamide: bisacrylamide (37.5:1) (Bio-Rad) gels. Electrophoresis was carried out using Protean II xi cells (Bio-Rad) for 18 hours at 390 volts, with a 12 % polyacrylamide gel concentration and run temperature of 13 ºC for amplicon I (part of intron 1, exon 2 and part of intron 2); for 16 hours at 390 volts, with 10 % polyacrylamide at 18 ºC for amplicon II (part of intron 2); and for 24 hours at 390 volts, with 10 % polyacrylamide plus 4 % glycerol at 15 ºC for amplicon III (part of intron 2 and exon 3). A silver staining technique described by Byun *et al.* (2009), was used to reveal the SSCP banding patterns.

4.1.8 Nucleotide sequencing

Based on the PCR-SSCP patterns identified, cattle that were homozygous with unique banding patterns for amplicons 1, 2 and 3 were subjected to direct sequencing at the Lincoln University DNA sequencing facility (Lincoln University, New Zealand). For these cattle, the amplicons were purified using a MiniElute™ PCR Purification Kit (Qiagen), and then directly sequenced using the original PCR primers to prime the sequencing reactions in both directions, using the forward and reverse primers.

For rare patterns that were only found in heterozygous form, the unique band was excised from the wet gel and incubated at 69 ºC for 1 hour as per Gong *et al.* (2011). A 1-µL aliquot of the incubated product was then mixed with 14-µL of PCR pre-mixture and then amplified using identical conditions to the original PCR reactions.

For each of the regions sequenced, the nucleotide sequences were aligned and other analyses undertaken using DNAMAN (Version 5.2.10, Lynnon BioSoft, Vaudreuil, Canada).

4.1.9 Statistical analysis

The Hardy-Weinberg equilibrium (HWE) for the leptin genotypes was analysed using an online chi-square calculator (http://www.oege.org/software/hwe-mr-calc.shtml). The statistics software IBM SPSS version 22 (IBM, NY, USA) was used to perform all other statistical analyses, and unless otherwise stated, significance was accepted when $P < 0.05$.

Of the three regions investigated, amplicon 3 was the most variable, and was used for association studies. Using General Linear Mixed-effects Models (GLMMs), associations between leptin variation and variation in milk component levels were tested. First, single-variant presence/absence models (fixed effects: DIM, age, herd and leptin variant) was used to ascertain which variant(s) should be analysed in subsequent multi-variant models. The multi-variant models included any variant that had a variant-FA trait association in the single-variant presence/absence analysis with a *P*-value of less than 0.200. This is a low threshold for the inclusion of a possibly explanatory factor in the model. The multi-variant models were again corrected for the fixed effects of DIM, age, herd and leptin variant, and with DIM and age fitted as covariates.

Next, the effect of genotypes on milk FA component levels, average bodyweights and milk traits were assessed using a GLMM (fixed effect: leptin genotype, DIM, age, herd) and multiple pair-wise comparisons with Bonferroni corrections.

However, the effect of sire of cow could not be included in the GLMMs, because some semen straws (sire genetics) used in New Zealand dairy cattle artificial insemination-based breeding approaches contain mixed-sire semen purchased from commercial semen producers. In these cases, it is impossible to ascertain individual sire identity. However, since the straws were mixed-semen straws and because different sires are used for different inseminations, in different years, it is unlikely that sire was a strongly confounding effect. Also, cow age and herd might be confounded with sire, but this cannot be confirmed.

4.2 Results

4.2.1 Nucleotide sequence variation

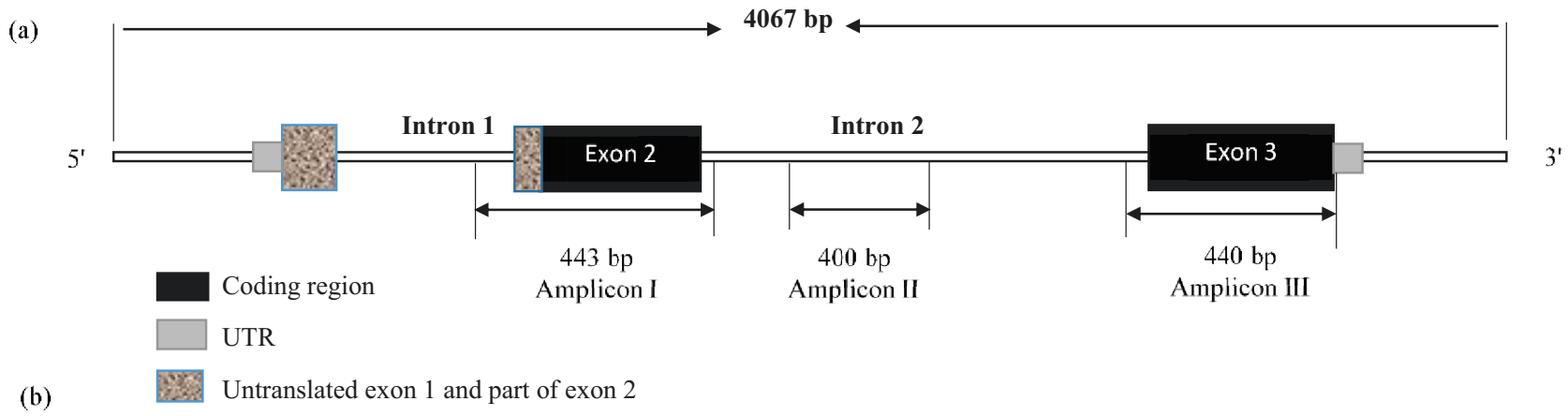
In this investigation, three pairs of primers were used to genotype 657 cattle from eight different Nigerian and New Zealand breeds. A total of three amplicons covering 1,152 bp of non-contiguous sequence were generated. The amplified regions included exons 2 and 3, and parts of introns 1 and 2 (Figure 4.1).

Three different PCR-SSCP analyses coupled with DNA sequencing were used to search for genetic variation in amplicons I, II and III in the cattle investigated. This technique revealed (Figure 4.2) four banding patterns (A_1, B_1, C_1 and D_1) in amplicon I (exon 2, and parts of introns 1 and 2), three patterns (A_2, B_2 and C_2) in amplicon II (intron 2 region) and three patterns (A_3, B_3 and C_3) in amplicon III (which spans exon 3 and the intron 2 boundary region). Of the four variants identified for amplicon I, the NZ HF × J cross-bred cattle had three variants (A_1, B_1, and C_1), and the Nigerian Red Bororo and cross-bred HF x WF cattle had three variants each (A_1, B_1, and D_1), while two variants (A_1 and B_1) were observed in the NZ Hereford, Angus, Shorthorn, Nigerian Sokoto Gudali, and White Fulani cattle (Table 4.1). In the amplicon II region, all the eight breeds were observed to carry the A_2 and B_2 variants, but only the Nigerian Red Bororo had the C_2 variant (Table 4.2). All eight breeds carried the A_3, B_3 and C_3 variants in the amplicon III region investigated (table 4.3).

A total of twelve nucleotide variations were identified across the three regions examined. Eleven were nucleotide substitutions; c.-107G/C and c.-130C/T located in intron 1, c.73C/T (p.Arg25Cys) in exon 2, c.144+808C/T, c.144+907C/T and c.144+908T/C identified in intron 2 and c.239C/T (p.Ala80Val), c.396C/T (p.Gly132), c.399T/C (p.Val133), c.411T/C (p.Ala137) and c.495C/T (p.Pro165) located in exon 3. The remaining one nucleotide variation was a deletion; c.144+42delC located in intron 2.

Of the twelve variations identified across the three regions examined, four nucleotide substitutions, which included a deletion (c.144+42delC) and three substitutions (c.144+808C/T, c.144+907C/T and c.144+908T/C) were novel and are reported here for the first time. The remaining eight variations, with their respective rs numberings have been previously reported in Ensembl (EMBL-EBI, Hinxton, United Kingdom, release 96 - April 2019 ARS-UCD1.2 for cow Leptin gene) table 4.4.

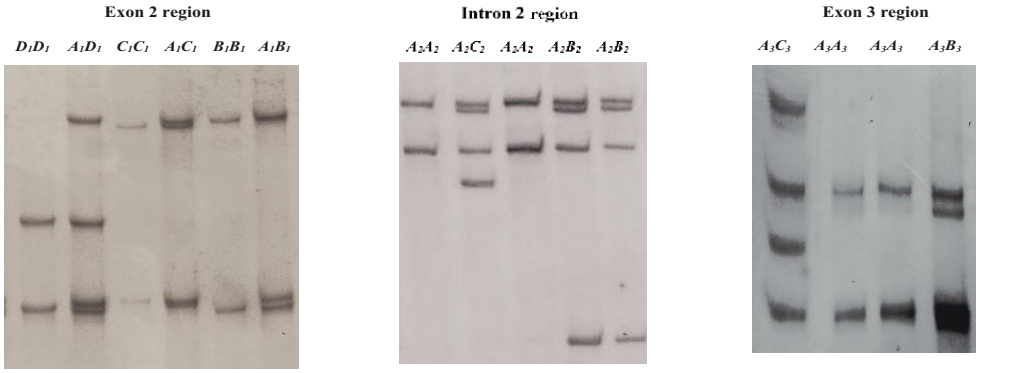
The alignment of the predicted amino acid sequence of leptin protein obtained from the four variants (A_1, B_1, C_1 and D_1) in exon 2 and three variants (A_3, B_3, and C_3) in exon 3 are given in Figure 4.3.



Amplicon	Primers 5'- 3'	Annealing temperature (ºC)	Polyacrylamide concentration	SSCP conditions	SSCP patterns	GenBank accession numbers
I	GTCTTTGAGGAGATGATAGCC GCTGTCTTTATGCCAGGGG	60	12%	390 V, 13 ºC, 18 hrs	4	MN082388, MN082389, MN082390, MN082391
II	AGCTAGTCAGGTTCCACAAG GGTTCTGCAAGGGTATTCAG	60	10%	390 V, 18 ºC, 16 hrs	3	MN069837, MN069838, MN069839
III	TTGCTCTCCCCTTCCTCCTG CTCAGGTTTCTTCCCTGGAC	60	10% + 4% Gly	390 V, 15 ºC, 24 hrs	3	MN119553, MN119554, MN119555

Figure 4.1 (a) Schematic representation of the bovine leptin gene (*LEP*) showing the amplified regions (based on GenBank accession number U50365.1).
Amplicon I comprises part of intron 1, the entirety of exon 2 and part of intron 2; amplicon II consists of part of intron 2 whereas amplicon III comprises of parts of Intron 2, entire exon 3 and part of 3'UTR. (b) Optimized PCR-SSCP conditions for each amplicon, banding patterns and accession numbers are shown.

(a)



(b)

	A_1	B_1	C_1	D_1
c.-130	C	T	T	C
c.-107	G	C	C	G
c.73	C	T	C	C
c.144+42	C	C	C	-

	A_2	B_2	C_2
c.144+808	C	T	C
c.144+907	C	C	T
c.144+908	T	T	C

	A_3	B_3	C_3
c.239	C	T	C
c.396	C	C	T
c.399	T	T	C
c.411	T	T	C
c.495	C	C	T

Figure 4.2 Nucleotide sequence variations obtained in the extended region of bovine leptin gene investigated
(a) PCR-SSCP banding patterns obtained in the three amplicons of bovine leptin gene investigated. (b) Nucleotide sequencing revealed the different nucleotide sequence variations identified in each of these amplicons.
A dash (-) represents a nucleotide deletion.

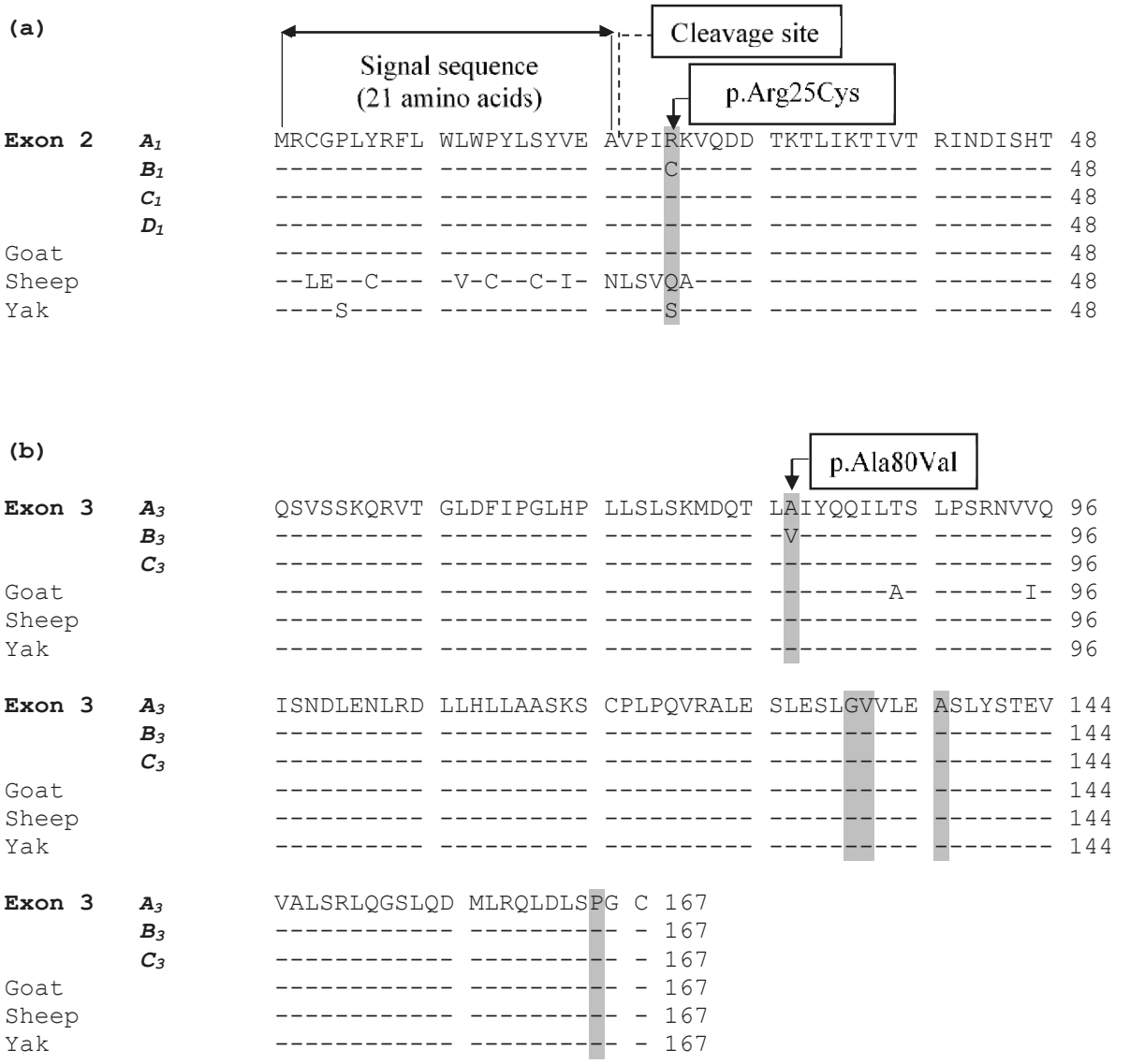


Figure 4.3 The amino acid sequence of bovine leptin gene obtained from the exon 2 and exon 3 regions examined in comparison to other species

(a) Shows the alignment of the predicted amino acid sequence of leptin protein obtained from the four variants (A_1, B_1, C_1 and D_1) identified in exon 2. The highlighted area in grey shows the position of amino acid change across the variant sequences and other species. (b) Alignment of the putative amino acid sequence of leptin protein from the three variants (A_3, B_3 and C_3) in exon 3, and from sequences in other species. Similarly, the highlighted area in grey shows the position of amino acid where the underlying nucleotide sequence is variable. The symbol '-'represents the same amino acid at the top (reference sequence). The amino acid sequence for cow was obtained from the coding sequences examined in this study, whereas that of goat was obtained from: GenBank: JQ739233.1, sheep: GenBank: MH716186.1 and yak: GenBank: EU603265.1.

Table 4.1 Variants and genotypes identified in each breed along with their respective frequencies in the exon 2-intron 2 region.

		Amplicon I			
Breed	**n**	**Variant**	**Variant Freq.%**	**Genotype**	**Genotype Freq.%**
HF × J cross	47	A_1	59.58	A_1A_1	31.91
		B_1	39.36	A_1B_1	53.19
		C_1	1.06	B_1B_1	12.77
				A_1C_1	2.13
Hereford	5	A_1	60.00	A_1A_1	40.00
		B_1	40.00	A_1B_1	40.00
				B_1B_1	20.00
Angus	5	A_1	60.00	A_1A_1	40.00
		B_1	40.00	A_1B_1	40.00
				B_1B_1	20.00
Shorthorn	4	A_1	75.00	A_1A_1	50.00
		B_1	25.00	A_1B_1	50.00
Sokoto Gudali	6	A_1	66.67	A_1A_1	33.33
		B_1	33.33	A_1B_1	66.66
White Fulani	12	A_1	66.67	A_1A_1	50.00
		B_1	33.33	A_1B_1	33.33
				B_1B_1	16.67
Red Bororo	12	A_1	58.33	A_1A_1	33.33
		B_1	25.00	A_1B_1	16.67
		D_1	16.67	B_1B_1	16.67
				A_1D_1	33.33
HF × WF cross	12	A_1	58.33	A_1A_1	33.33
		B_1	25.00	A_1B_1	16.67
		D_1	16.67	B_1B_1	16.67
				A_1D_1	33.33

Table 4.2 Variants and genotypes identified in each breed along with their respective frequencies in the intron 2 region in bovine.

		Amplicon II			
Breed	**n**	**Variant**	**Variant Freq.%**	**Genotype**	**Genotype Freq.%**
HF × J cross	47	A_2	92.55	A_2A_2	87.23
		B_2	7.45	A_2B_2	10.64
				B_2B_2	2.13
Hereford	10	A_2	60.00	A_2A_2	40.00
		B_2	40.00	A_2B_2	40.00
				B_2B_2	20.00
Angus	10	A_2	60.00	A_2A_2	40.00
		B_2	40.00	A_2B_2	40.00
				B_2B_2	20.00
Shorthorn	6	A_2	83.33	A_2A_2	66.67
		B_2	16.67	A_2B_2	33.33
Sokoto Gudali	6	A_2	83.33	A_2A_2	66.67
		B_2	16.67	A_2B_2	33.33
White Fulani	14	A_2	85.71	A_2A_2	71.43
		B_2	14.29	A_2B_2	28.57
Red Bororo	16	A_2	68.75	A_2A_2	37.50
		B_2	12.50	A_2B_2	25.00
		C_2	18.75	A_2C_2	37.50
HF × WF cross	6	A_2	66.67	A_2A_2	50.00
		B_2	33.33	A_2B_2	33.33
				B_2B_2	16.67

Table 4.3 Variants and genotypes identified in each of the eight breeds investigated, along with their respective frequencies in the entire exon 3 and parts of intron 2 splice donor site.

		Amplicon III			
Breed	**n**	**Variant**	**Variant Freq.%**	**Genotype**	**Genotype Freq.%**
HF × J cross	72	A_3	66.67	A_3A_3	43.06
		B_3	23.61	A_3B_3	47.22
		C_3	9.72	A_3C_3	9.72
Hereford	8	A_3	62.50	A_3A_3	25.00
		B_3	12.50	A_3B_3	25.00
		C_3	25.00	A_3C_3	50.00
Angus	8	A_3	62.50	A_3A_3	25.00
		B_3	18.75	A_3B_3	37.50
		C_3	18.75	A_3C_3	37.50
Shorthorn	8	A_3	62.50	A_3A_3	25.00
		B_3	18.75	A_3B_3	37.50
		C_3	18.75	A_3C_3	37.50
Sokoto Gudali	6	A_3	58.33	A_3A_3	33.33
		B_3	16.67	A_3B_3	16.67
		C_3	25.00	A_3C_3	50.00
White Fulani	6	A_3	66.67	A_3A_3	33.33
		B_3	25.00	A_3B_3	50.00
		C_3	8.33	A_3C_3	16.67
Red Bororo	6	A_3	58.33	A_3A_3	33.33
		B_3	16.67	A_3B_3	16.67
		C_3	25.00	A_3C_3	50.00
HF × WF cross	6	A_3	58.33	A_3A_3	16.67
		B_3	25.00	A_3B_3	33.33
		C_3	16.67	A_3C_3	50.00

Table 4.4 Nucleotide variations across three regions of bovine leptin gene in eight breeds and crosses of New Zealand and Nigerian cattle

	Amplicon I[2]				Amplicon II			Amplicon III				
Breed[1]	c.-130 (rs29004484)	c.-107 (rs29004485)	c.73 (Arg25Cys) (rs29004488)	[3]c.144+ 42delC	c.144+ 808	c.144+ 907	c.144+ 908	c.239 (p.Ala80Val) (rs29004508)	c.396 (p.Gly132) (rs29004509)	c.399 (p.Val133) (rs29004510)	c.411 (p.Ala137) (rs29004511)	c.495 (p.Pro165) (rs29004512)
Hereford	C/T	G/C	C/T	C	C/T	C	T	C/T	C/T	T/C	T/C	C/T
Angus	C/T	G/C	C/T	C	C/T	C	T	C/T	C/T	T/C	T/C	C/T
Shorthorn	C/T	G/C	C/T	C	C/T	C	T	C/T	C/T	T/C	T/C	C/T
HF × J-cross	C/T	G/C	C/T	C	C/T	C	T	C/T	C/T	T/C	T/C	C/T
SG	C/T	G/C	C/T	C	C/T	C	T	C/T	C/T	T/C	T/C	C/T
RB	C/T	G/C	C/T	C/-	C/T	C/T	T/C	C/T	C/T	T/C	T/C	C/T
WF	C/T	G/C	C/T	C	C/T	C	T	C/T	C/T	T/C	T/C	C/T
HF × WF-cross	C/T	G/C	C/T	C/-	C/T	C	T	C/T	C/T	T/C	T/C	C/T

[1]NZ HF × J-cross = Holstein Friesian × Jersey-cross, SG = Sokoto Gudali, RB = Red Bororo, WF = White Fulani and WF × HF = White Fulani × Holstein Friesian-cross.

[2]Amplicon I (Intron 1/exon 2/intron 2), amplicon II (Intron 2) amplicon III (Intron 2/Exon) regions.

[3] The symbol (-) represents nucleotide deletion.

4.2.2 Single variant models

Association studies were carried out between the leptin variants identified in exon 3 (amplicon III) and milk FA composition, milk production traits and bodyweight in NZ HF × J-cross cows. The results of the general linear mixed effect models revealed that the presence (or absence) of leptin A_3, B_3 and C_3 in a cow's genotype was associated with the quantity of milk FA methyl ester (FAME). The different variants had varying effects on the composition of milk FA levels as detailed in table 4.5. The presence of variant A_3 (the most common variant) was associated with decreased C15:1, C18:1 *trans*-11, C18:1 *trans*-9, *cis*-12, C18:1 all *trans*, C22:0 and C24:0 ($P < 0.05$) levels. The presence of variant B_3 was associated with decreased C6:0 level, C11:0 level, and C20:0 level, but increased C10:1 index, C24:0 and C17:0 *iso* ($P < 0.05$) levels. Variant C_3 was associated with decreased C13:0 *anteiso* ($P < 0.05$) level, and was found to also be associated with decreased MCFA ($P < 0.05$) level, but only when B_3 was included in the model.

The presence of variant B_3 was associated with decreased bodyweight in the single variant model ($P < 0.05$), even though this effect was lost when other variants were included in the model (table 4.7), and when the genotype model was analysed. Although there was no statistically significant result in the genotype model for bodyweight, an interesting trend was noted as cows carrying the A_3B_3 genotype had lower average bodyweights ($P = 0.099$). The A_3B_3 genotype appeared to have a more pronounced effect in decreasing average bodyweight (494.09 ± 5.74), compared to the A_3A_3 (504.09 ± 6.72) and the A_3C_3 (500.91 ± 7.12) genotypes (table 4.8).

4.2.3 Genotype model

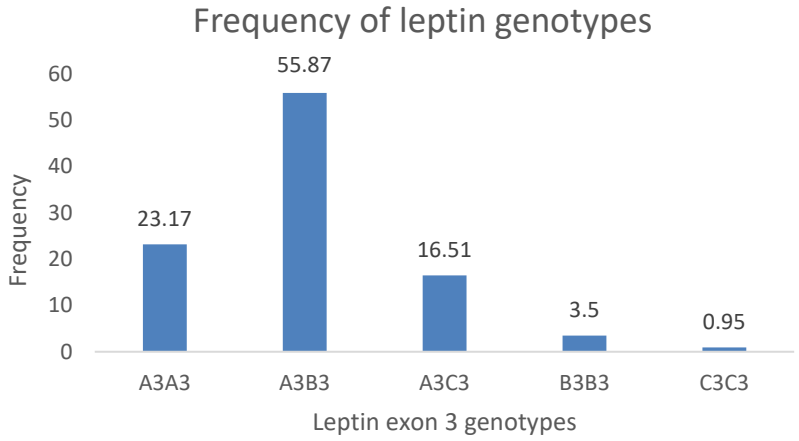


Figure 4.4 Leptin genotypes and their respective frequencies.
Cows with leptin genotypes A_3A_3 (n = 73), A_3B_3 (n = 176) and A_3C_3 (n = 52) with a frequency greater than 5% were analysed for milk FA composition and average bodyweight. Other genotypes; B_3B_3 (n = 11) and C_3C_3 (n = 3) had frequencies below 5% and were not included in this model.

The three different genotypes (A_3A_3, A_3B_3 and A_3C_3) analysed were found to have varying effects on the composition of milk FA levels, which were consistent with the results of the single variant model. Cows carrying the A_3A_3 (most common) genotype had higher levels of SFA, but when one copy of the A_3 variant is replaced by B_3 or C_3 variant, the resulting heterozygous genotype (A_3B_3 or A_3C_3) changes the levels of SFA in milk. Cows carrying the A_3B_3 genotype had lower marginal mean levels of the SCFA C8:0, grouped MCFA and the individual MCFAs C10:0, C11:0 and C13:0, but an increased level of C24:0 ($P < 0.05$). Genotype A_3C_3 was associated with decreased MCFA and the individual levels of C10:0, C11:0 and C13:0 ($P < 0.05$) (Table 4.6).

There was no association revealed between *LEP* genotypes and milk yield and fat percentage. However, cows carrying the A_3B_3 genotype had a decreased percentage of protein in their milk ($P < 0.05$) (Table 4.9).

Table 4.5 Associations between bovine leptin variants and the quantity of individual and grouped milk FA methyl ester (FAME) in NZ HF × J-cross cows' milk

			mean ± SE[1]				
	Variants	**Other Variants in model**	**Absent**	**n**	**Present**	**n**	***p***
C6:0	*A3*	none	1.554 ± 0.035	14	1.559 ± 0.017	301	0.871
	B3	none	1.575 ± 0.019	128	1.549 ± 0.018	187	**0.046**
	C3	none	1.559 ± 0.017	260	1.557 ± 0.022	55	0.938
C8:0	*A3*	none	1.169 ± 0.029	14	1.183 ± 0.014	301	0.588
	B3	none	1.196 ± 0.016	128	1.176 ± 0.015	187	*0.066*
	C3	none	1.185 ± 0.015	260	1.174 ± 0.018	55	0.425
C10:0	*A3*	none	3.134 ± 0.112	14	3.243 ± 0.055	301	0.273
	B3	none	3.284 ± 0.061	128	3.218 ± 0.057	187	*0.107*
	C3	none	3.253 ± 0.056	260	3.184 ± 0.071	55	0.202
C10:1	*A3*	none	0.270 ± 0.014	14	0.284 ± 0.007	301	0.250
	B3	none	0.280 ± 0.007	128	0.286 ± 0.007	187	0.268
	C3	none	0.285 ± 0.007	260	0.274 ± 0.009	55	*0.093*
C11:0	*A3*	none	0.055 ± 0.006	14	0.057 ± 0.003	301	0.623
	B3	none	0.060 ± 0.003	128	0.056 ± 0.003	187	**0.027**
	C3	none	0.058 ± 0.003	260	0.055 ± 0.004	55	0.354
C12:0	*A3*	none	3.746 ± 0.151	14	3.945 ± 0.074	301	*0.138*
	B3	none	3.980 ± 0.082	128	3.920 ± 0.076	187	0.280
	C3	none	3.960 ± 0.075	260	3.853 ± 0.095	55	*0.138*
	A3	*C3*	3.728 ± 0.156	14	3.926 ± 0.085	301	*0.139*
	C3	*A3*	3.908 ± 0.114	260	3.802 ± 0.128	55	*0.140*
C12:1	*A3*	none	0.082 ± 0.006	14	0.091 ± 0.003	301	*0.057*
	B3	none	0.091 ± 0.003	128	0.091 ± 0.003	187	0.796
	C3	none	0.092 ± 0.003	260	0.087 ± 0.004	55	*0.092*
	A3	*C3*	0.081 ± 0.006	14	0.090 ± 0.003	301	*0.058*
	C3	*A3*	0.089 ± 0.005	260	0.084 ± 0.006	55	*0.094*
C13:0 *antiso*	*A3*	none	0.038 ± 0.001	14	0.038 ± 0.001	301	0.616
	B3	none	0.037 ± 0.001	128	0.038 ± 0.001	187	*0.331*
	C3	none	0.038 ± 0.001	260	0.037 ± 0.001	55	**0.025**
C13:0	*A3*	none	0.118 ± 0.008	14	0.117 ± 0.004	301	0.856
	B3	none	0.120 ± 0.004	128	0.115 ± 0.004	187	*0.103*
	C3	none	0.118 ± 0.004	260	0.113 ± 0.005	55	0.229
C14:0	*A3*	none	1.563 ± 0.258	14	12.560 ± 0.126	301	0.993
	B3	none	12.641 ± 0.139	128	12.516 ± 0.130	187	*0.185*
	C3	none	12.581 ± 0.127	260	12.461 ± 0.162	55	0.334
C14:1	*A3*	none	0.892 ± 0.067	14	0.955 ± 0.032	301	0.285
	B3	none	0.929 ± 0.036	128	0.968 ± 0.033	187	*0.103*
	C3	none	0.959 ± 0.033	260	0.930 ± 0.042	55	0.353
C15:0	*A3*	none	1.489 ± 0.055	14	1.459 ± 0.027	301	0.548
	B3	none	1.477 ± 0.030	128	1.450 ± 0.028	187	*0.186*
	C3	none	1.462 ± 0.027	260	1.448 ± 0.035	55	0.607

(Continue on next page)

(Continued)

C15:1	A_3	none	0.301 ± 0.010	14	0.282 ± 0.005	301	**0.028**
	B_3	none	0.280 ± 0.005	128	0.283 ± 0.005	187	0.541
	C_3	none	0.283 ± 0.005	260	0.277 ± 0.006	55	0.258
C17:0 *iso*	A_3	none	0.570 ± 0.021	14	0.557 ± 0.010	301	0.498
	B_3	none	0.546 ± 0.011	128	0.564 ± 0.010	187	**0.014**
	C_3	none	0.562 ± 0.010	260	0.539 ± 0.013	55	**0.020**
	B_3	C_3	0.546 ± 0.012	128	0.562 ± 0.011	187	**0.034**
	C_3	B_3	0.560 ± 0.011	260	0.542 ± 0.014	55	*0.090*
C17:0	A_3	none	0.860 ± 0.026	14	0.872 ± 0.013	301	0.581
	B_3	none	0.880 ± 0.014	128	0.868 ± 0.013	187	*0.184*
	C_3	none	0.871 ± 0.013	260	0.876 ± 0.016	55	0.677
C18:1 *trans*-11	A_3	none	3.160 ± 0.229	14	2.707 ± 0.112	301	**0.026**
	B_3	none	2.729 ± 0.125	128	2.707 ± 0.116	187	0.791
	C_3	none	2.722 ± 0.114	260	2.681 ± 0.145	55	0.708
C18:1 *cis*-9, *trans*-12	A_3	none	0.083 ± 0.007	14	0.074 ± 0.004	301	0.161
	B_3	none	0.076 ± 0.004	128	0.074 ± 0.004	187	0.407
	C_3	none	0.074 ± 0.004	260	0.075 ± 0.005	55	0.899
C18:1 *trans*-9, *cis*-12	A_3	none	0.564 ± 0.037	14	0.488 ± 0.018	301	**0.020**
	B_3	none	0.494 ± 0.020	128	0.486 ± 0.019	187	0.546
	C_3	none	0.489 ± 0.018	260	0.487 ± 0.023	55	0.912
C18:3 *cis*-9, 12, 15	A_3	none	0.720 ± 0.034	14	0.757 ± 0.017	301	0.215
	B_3	none	0.746 ± 0.019	128	0.763 ± 0.017	187	*0.168*
	C_3	none	0.759 ± 0.017	260	0.747 ± 0.022	55	0.490
C20:0	A_3	none	0.125 ± 0.005	14	0.125 ± 0.003	301	0.960
	B_3	none	0.128 ± 0.003	128	0.123 ± 0.003	187	**0.028**
	C_3	none	0.124 ± 0.003	260	0.129 ± 0.003	55	*0.059*
C20:1 *cis*-11	A_3	none	0.068 ± 0.004	14	0.074 ± 0.002	301	*0.078*
	B_3	none	0.075 ± 0.002	128	0.074 ± 0.002	187	0.683
	C_3	none	0.074 ± 0.002	260	0.075 ± 0.003	55	0.782
C20:4 *cis*-5, 8, 11, 14	A_3	none	0.036 ± 0.002	14	0.034 ± 0.001	301	*0.182*
	B_3	none	0.034 ± 0.001	128	0.034 ± 0.001	187	0.819
	C_3	none	0.034 ± 0.001	260	0.033 ± 0.001	55	0.289
C22:0	A_3	none	0.074 ± 0.004	14	0.064 ± 0.002	301	**0.006**
	B_3	none	0.062 ± 0.002	128	0.065 ± 0.002	187	*0.056*
	C_3	none	0.064 ± 0.002	260	0.064 ± 0.003	55	0.772
	A_3	B_3	0.073 ± 0.004	14	0.064 ± 0.002	301	**0.009**
	B_3	A_3	0.066 ± 0.005	128	0.069 ± 0.005	187	*0.093*
C24:0	A_3	none	0.051 ± 0.003	14	0.044 ± 0.001	301	**0.009**
	B_3	none	0.043 ± 0.002	128	0.045 ± 0.001	187	**0.024**
	C_3	none	0.044 ± 0.001	260	0.045 ± 0.002	55	0.842
C22:5 *cis*-9, 10, 13, 16,19	A_3	none	0.124 ± 0.007	14	0.117 ± 0.004	301	0.315
	B_3	none	0.117 ± 0.004	128	0.117 ± 0.004	187	0.919
	C_3	none	0.118 ± 0.004	260	0.113 ± 0.005	55	*0.114*

(Continue on next page)

(Continued)

MCFA	A_3	none	20.611 ± 0.498	14	20.932 ± 0.243	301	0.469
	B_3	none	21.101 ± 0.269	128	20.830 ± 0.251	187	*0.137*
	C_3	none	20.979 ± 0.246	260	20.672 ± 0.313	55	*0.199*
	B_3	C_3	21.066 ± 0.427	128	20.536 ± 0.434	187	**0.013**
	C_3	B_3	21.112 ± 0.366	260	20.446 ± 0.419	55	**0.017**
C18:1 all *trans*	A_3	none	3.460 ± 0.234	14	2.993 ± 0.114	301	**0.025**
	B_3	none	3.017 ± 0128	128	2.993 ± 0.119	187	0.781
	C_3	none	3.007 ± 0.117	260	2.971 ± 0.148	55	0.751
all C18:3	A_3	none	0.794 ± 0.035	14	0.831 ± 0.017	301	0.233
	B_3	none	0.819 ± 0.019	128	0.837 ± 0.018	187	*0.164*
	C_3	none	0.833 ± 0.017	260	0.822 ± 0.022	55	0.511
Omega 3	A_3	none	0.937 ± 0.035	14	0.962 ± 0.017	301	0.416
	B_3	none	0.950 ± 0.019	128	0.968 ± 0.018	187	*0.172*
	C_3	none	0.965 ± 0.017	260	0.947 ± 0.022	55	0.307
C10:1 index	A_3	none	7.948 ± 0.427	14	8.136 ± 0.209	301	0.621
	B_3	none	7.921 ± 0.230	128	8.249 ± 0.214	187	**0.036**
	C_3	none	8.160 ± 0.211	260	7.996 ± 0.269	55	0.422
C12:1 index	A_3	none	2.110 ± 0.111	14	2.262 ± 0.054	301	*0.128*
	B_3	none	2.219 ± 0.060	128	2.281 ± 0.056	187	*0.126*
	C_3	none	2.271 ± 0.055	260	2.201 ± 0.070	55	*0.194*
	A_3	B_3, C_3	2.091 ± 0.116	14	2.255 ± 0.062	301	*0.099*
	B_3	A_3, C_3	2.161 ± 0.101	128	2.229 ± 0.097	187	*0.098*
	C_3	A_3, B_3	2.225 ± 0.089	260	2.163 ± 0.099	55	0.267
C14:1 index	A_3	none	6.608 ± 0.483	14	7.072 ± 0.236	301	0.281
	B_3	none	6.843 ± 0.261	128	7.185 ± 0.243	187	*0.054*
	C_3	none	7.090 ± 0.239	260	6.937 ± 0.305	55	0.510

[1]Predicted means and standard error of those means derived from GLMM. Cow age, herd, DIM, and leptin variants were fitted to the models as fixed effects, with DIM and age fitted as covariates. $0.05 < P < 0.2$ in italics and $P < 0.05$ in bold.

Table 4.6 Associations between milk FA levels and leptin genotypes

	Mean ± SE[1] (g/100 g milk FA)			*p*
	A₃A₃ n = 73	***A₃B₃*** n = 176	***A₃C₃*** n = 52	
C6:0	1.589 ± 0.021	1.552 ± 0.018	1.557 ± 0.022	0.056
C8:0	1.213 ± 0.018[a]	1.178 ± 0.015[b]	1.176 ± 0.019[ab]	**0.019**
C10:0	3.357 ± 0.067[a]	3.223 ± 0.057[b]	3.196 ± 0.072[b]	**0.014**
C10:1	0.285 ± 0.008	0.287 ± 0.007	0.277 ± 0.009	0.340
C11:0	0.064 ± 0.003[a]	0.056 ± 0.003[b]	0.056 ± 0.004[b]	**0.003**
C12:0	4.072 ± 0.091	3.928 ± 0.077	3.873 ± 0.096	**0.044**
C12:1	0.094 ± 0.003	0.092 ± 0.003	0.089 ± 0.004	0.317
C13:0 *antiso*	0.038 ± 0.001	0.038 ± 0.001	0.037 ± 0.001	0.087
C13:0	0.126 ± 0.005[a]	0.115 ± 0.004[b]	0.114 ± 0.005[b]	**0.010**
C14:0	12.766 ± 0.155	12.502 ± 0.131	12.491 ± 0.164	0.057
C14:1	0.929 ± 0.040	0.973 ± 0.034	0.937 ± 0.042	0.269
C15:0	1.499 ± 0.033	1.446 ± 0.028	1.458 ± 0.035	0.100
C15:1	0.284 ± 0.006	0.282 ± 0.005	0.278 ± 0.006	0.604
C17:0 *iso*	0.553 ± 0.012	0.564 ± 0.010	0.541 ± 0.013	0.077
C17:0	0.881 ± 0.016	0.866 ± 0.013	0.876 ± 0.016	0.408
C18:1 *trans*-11	2.796 ± 0.138	2.697 ± 0.117	2.704 ± 0.147	0.613
C18:1 *cis*-9, *trans*-12	0.077 ± 0.004	0.074 ± 0.004	0.075 ± 0.005	0.534
C18:1 *trans*-9, *cis*-12	0.506 ± 0.022	0.485 ± 0.019	0.493 ± 0.024	0.465
C18:3 *cis*-9, 12, 15	0.745 ± 0.021	0.767 ± 0.018	0.744 ± 0.022	0.226
C20:0	0.126 ± 0.003	0.123 ± 0.003	0.129 ± 0.003	0.091
C20:1 *cis*-11	0.074 ± 0.003	0.074 ± 0.002	0.075 ± 0.003	0.991
C20:4 *cis*-5, 8, 11, 14	0.035 ± 0.001	0.034 ± 0.001	0.033 ± 0.001	0.289
C22:0	0.061 ± 0.002	0.065 ± 0.002	0.063 ± 0.003	0.073
C24:0	0.042 ± 0.002[a]	0.045 ± 0.001[b]	0.044 ± 0.002[ab]	**0.034**
C22:5 *cis*-9, 10, 13, 16, 19	0.121 ± 0.004	0.118 ± 0.003	0.113 ± 0.004	0.122
MCFA	21.408 ± 0.299[a]	20.830 ± 0.253[b]	20.735 ± 0.317[b]	**0.019**
C18:1 all *trans*	3.081 ± 0.142	2.982 ± 0.120	2.995 ± 0.150	0.632
all C18:3	0.818 ± 0.021	0.841 ± 0.018	0.819 ± 0.023	0.235
Omega 3	0.954 ± 0.021	0.972 ± 0.018	0.945 ± 0.022	0.220
C10:1 index	7.888 ± 0.257	8.276 ± 0.218	8.030 ± 0.273	0.104
C12:1 index	2.236 ± 0.067	2.292 ± 0.057	2.221 ± 0.071	0.319
C14:1 index	6.787 ± 0.291	7.222 ± 0.246	6.974 ± 0.308	0.115

[1]Predicted means and standard error of those means derived from GLMM. Cow age, herd, DIM, and leptin genotypes were fitted to the models as fixed effects, with DIM and age fitted as covariates. $P < 0.05$ in bold, and means within a row that do not share a superscript letter are separated by Bonferroni test at $P < 0.05$.

Table 4.7 Associations between leptin variants and average bodyweights

			mean ± SE [1]					
	Variant	Other Variants in model	Absent		n	Present	n	*p*
weight	*A3*	none	481.03	± 11.32	13	496.99 ± 5.47	299	*0.111*
	B3	none	501.99	± 6.02	128	493.93 ± 5.65	184	**0.045**
	C3	none	496.54	± 5.56	257	498.55 ± 7.01	55	0.704
	A3	*B3*	483.01	± 11.78	13	497.70 ± 6.31	299	*0.142*
	B3	*A3*	498.29	± 8.65	128	490.51 ± 8.28	184	*0.054*

1Predicted means and standard error of those means derived from GLMM. Cow age, herd, DIM, and leptin variants were fitted to the models as fixed effects, with DIM and age fitted as covariates. $0.05 < P < 0.2$ in italics and $P < 0.05$ in bold.

Table 4.8 Associations between leptin genotypes and marginal mean bodyweights

	A3A3	*A3B3*	*A3C3*	***p***
	n = 73	n = 176	n = 52	
weight	504.09 ± 6.72	494.09 ± 5.74	500.91 ± 7.12	*0.099*

1Predicted means and standard error of those means derived from GLMM. Cow age, herd, DIM, and leptin genotypes were fitted to the models as fixed effects, with DIM and age fitted as covariates.

Table 4.9 Association between milk production traits and leptin genotypes in NZ HF × J-cross cows

	Mean ± SE[1] (g/100 g milk FA)			
Milk Production traits	*A3A3*	*A3B3*	*A3C3*	***p***
	n = 110	n = 249	n =84	
Milk Yield (L/day)	22.730 ± 0.318	22.694 ± 0.213	22.538 ± 0.366	0.915
Fat %	5.024 ± 0.051	4.929 ± 0.034	5.041 ± 0.058	0.137
Protein %	4.081 ± 0.026^{a}	3.996 ± 0.017^{b}	4.016 ± 0.030^{ab}	**0.024**

[1]Predicted means and standard error of those means derived from GLMM. Cow age, DIM, herd, and leptin genotype were fitted to the models as fixed effects, with DIM and age fitted as covariates. $P < 0.05$ in bold and means within a row that do not share a superscript letter are separated by Bonferroni test at $P < 0.05$.

4.3 Discussion

4.3.1 Nucleotide sequence variation in LEP

Of the twelve variations identified in this study, four are novel, and eight have been described previously. Of the eight reported previously, two would if expressed, change the amino acid sequence of the protein. These are c.73C/T (p.Arg25Cys) in exon 2 and c.239C/T (p.Ala80Val) in exon 3. The p.Arg25Cys substitution would result from the C/T substitution located 73 bp from the start of exon 2 in a region that is proximal to the signal sequence of leptin and that has been described previously (Konfortov *et al.*, 1999; Buchanan *et al.*, 2002; Lagonigro *et al.*, 2003). The signal sequence contains 21 amino acids and it is cleaved off before leptin is excreted from adipose tissue (Zhang *et al.*, 1997). The predicted cleavage site of the signal sequence is C terminal to an alanine at position 21 (Zhang *et al.*, 1997). The p.Ala80Val would result from a C/T substitution located 95 bp from the start of exon 3. It has also been described previously (Konfortov *et al.*, 1999; Haegeman *et al.*, 2000).

Zhang *et al.* (1997) suggested that the p.Arg25Cys change is situated in the first of the four alpha-helices of the leptin protein, and at a location that varies considerably between species. They reported glutamine in primates and mice, arginine in dogs, tryptophan in pigs and histidine in rats. Therefore, it is likely that the p.Arg25Cys change is allowed because it does not have functional significance. This view was also shared by Konfortov *et al.* (1999). It is also conceivable that this position may be polymorphic in other species. This is supported by the amino acid alignment of leptin protein from different ruminant species (Figure 4.3), with position 25 of the protein in goats being an arginine residue, while sheep have glutamine and yak have serine. Variation at this position therefore appears to be tolerated. Since arginine and cysteine have alkaline and neutral side chains respectively, this variation could be considered a non-conservative substitution.

However, Buchanan *et al.* (2002) hypothesized that the p.Arg25Cys change has a functional effect on the leptin molecule. They suggested that the presence of cysteine in the α-helix of the leptin molecule may disrupt the binding of leptin to its receptor. This is supported by the observation that the leptin receptor contains a conserved trough typical of haemopoietic cytokine receptors, into which the A and D helixes of haemopoietic cytokines dock (Sprang *et al.*, 1993). They argued that a change between two very different amino acids, such as arginine and cysteine at this location, may have deleterious effect on this process. Also, their report suggests that the presence of another unpaired cysteine in the leptin molecule could destabilize the disulphide bridge found between the two existing cysteines, thus affecting the structure of the leptin molecule, and consequently, its function (Rock *et al.*, 1996). Studies have shown that this disulphide bridge is critical for biological function (Rock *et al.*, 1996; Zhang *et al.*, 1997).

In comparison, the p.Ala80Val polymorphism is located in a conserved region of the leptin protein (Komisarek *et al.*, 2005). The c.396C/T (p.Gly132), c.399T/C (p.Val133), c.411T/C (p.Ala137) and

c.495C/T (p.Pro165) nucleotide sequence differences in exon 3 are all synonymous, and occur in a conserved region of leptin.

Four nucleotide changes, which included a deletion (c.144+42delC) and three substitutions (c.144+808C/T, c.144+907C/T and c.144+908T/C) occurred in the intron 2 region. They are reported here for the first time. In amplicon I, the D_1 variant, carrying the deletion c.144+42delC was identified only in the Nigerian Red Bororo and the cross-bred Holstein Friesian × White Fulani cattle breeds. In amplicon II, the B_2 variant, with the haplotype c.144+808C, c.144+907T and c.144+908C was found in all the New Zealand and Nigerian breeds investigated, whereas the C_2 variant for amplicon II, which carry's the substitutions c.144+808T, c.144+907C and c.144+908T was only detected in the Red Bororo breed. The occurrence of variant B_2 with GenBank accession number MN082389 in all the nine breeds (New Zealand and Nigerian cattle breeds) investigated, suggests that this variation is shared across cattle of *Bos taurus* and *Bos indicus* (African Zebu cattle) origin. This result is not surprising, given the close relationship between *Bos taurus* and *Bos indicus* in terms of origin. There are suggestions that these two species originated from a common ancestor, and that the auroch (*Bos primigenius*) is the progenitor of all taurine and zebu (African *Bos indicus*) cattle (Edwards *et al.*, 2007). Estimates of divergence times for *Bos taurus* and *Bos indicus* from a common ancestor are all pre-Neolithic and range from approximately 2 million to 330,000 years ago, depending on the genetic markers and the calibration of the evolutionary molecular clock (Hiendleder *et al.*, 2008).

There are also some reports that suggest the *Bos indicus* populations may have been produced at a later date through breeding and selection from *Bos taurus* cattle (Epstein 1973; Epstein and Mason 1984). Either way, each of these aforementioned studies highlighted the evolutionary link between *Bos taurus* and *Bos indicus* cattle. The sharing of common haplotypes suggest that cattle studied here might be more closely related, although it does not provide definitive proof. The suggestion is supported by Dunner *et al.* (2003) who suggested that the pattern of haplotype sharing is an indicator of the history of the different bovine breeds, and thus the distribution of shared haplotypes is useful in describing population relationships.

Some reports based on archaeological and genetic evidence, contend that domesticated zebu populations were developed independently to *Bos taurus* by a separate group of early pastoralists, and that the southern Asian subspecies of aurochs (*Bos primigenius namadicus*) is most likely the progenitor of domesticated zebu cattle (Meadow 1993; Loftus *et al.*, 1994a; Bradley *et al.*, 1996). However, analysis of mitochondrial DNA (mtDNA) sequence variation in extant cattle populations, Loftus *et al.* (1994a) suggests that the taurine and zebu lineages probably diverged at least 210,000 years ago, which probably places them outside the range required for the human-mediated development of zebu cattle from taurine progenitors.

In the amplicon I region investigated, the D_1 variant carrying the novel deletion c.144+42delC, was proximal to the putative splice-donor site. The deletion variant was only identified in the Nigerian

Red Bororo cattle and the cross-bred Holstein Friesian × White Fulani-cross cattle. This suggests that this variation may only be found in the African Zebu cattle, since it wasn't observed in the NZ HF × J-cross cattle or the other New Zealand beef breeds. The occurrence of this variation around the exon/intron boundary may possibly have functional implications for gene expression and protein assembly.

To illustrate this point, Sjakste *et al.* (2011) investigated variation in *MSTN* in Latvian Darkhead sheep and reported nucleotide sequence variations (c.373+18G/T and c.373+101C/T) at the splice-donor site and showed that the G nucleotide in the c.373+18 position, initiated single-strandedness in the first CUG repeat and G triplet location. They further suggested that hairpin-loop development in that region could be followed by rearrangement of the spatial topology between the two other CUG repeats. Such perturbation of the pre-mRNA secondary structure could potentially influence sequence interaction with different regulatory proteins and the efficiency of lariat formation, and this could then potentially affect transcription and splicing efficiency (Chasin, 2007; Hiller *et al.*, 2007; Aznarez *et al.*, 2008). There is also evidence suggesting that introns are functionally active participants in gene and genome functionality, as they can encode regulatory elements (Yutzey *et al.*, 1989) that participate in splicing, transcription, and recombination events. It will therefore be interesting to further investigate the effect of this deletion in the Red Bororo and White Fulani × Holstein Friesian-cross cattle.

In the amplicon I region examined, the C_1 variant (carrying the nucleotide sequence variations c.-107C and c.-130T) had the lowest frequency, which was also reflected at the genotype level. Again, this is not surprising, especially considering the fact that this variant was only identified in one heterozygous sample (A_1C_1), out of the 47 samples of NZ HF × J-cross cows investigated. Since this variant had very low frequency in the cattle studied here, it may in future be difficult to ascertain how it affects the function of the gene. The investigation of more cows and more breeds may resolve this issue, especially as in the Ensembl report on Iranian *Bos taurus* from the NextGen Project (The European Union's Seventh Framework Programme [FP7/2010-2014]), suggests the G and C nucleotides in c.-107G/C (rs29004485) had a frequency of 69% and 31% respectively [remapped to ARS-UCD1.2] (release 150), and the C and T alleles in the nucleotide variation c.-130C/T (rs29004484) had a frequency of 69% and 31% respectively.

4.3.2 Associations between leptin variants in exon 3 and milk production traits, milk FA composition and average bodyweight in NZ HF × J – cross cows

There has been several studies demonstrating associations between leptin gene polymorphisms and economically important traits in cattle. However, this is the first study to investigate the effect of leptin gene variations in exon 3 with bodyweight, milk production traits and the composition of milk FAs in NZ HF × J-cross cows.

Overall, the results presented in this chapter revealed that the presence of variant B_3 was associated with a lower adjusted mean level of some short and medium-chain length SFA levels, but higher levels of some long-chain FA, branched FA as well as MUFA index. Variant C_3 was found to be associated with a decreased branched SFA. These results were also consistent in the genotype model, as cows carrying the A_3B_3 genotype showed lower levels of short and medium-chain length SFAs, but higher levels of LCFA, while A_3C_3 was associated with decreased medium-chain length SFA levels. There was no associations between bovine leptin genotypes and average bodyweight, milk yield and milk fat percentage, but an association was observed between leptin genotypes and decreased percentage of protein in milk.

4.3.3 Effects of leptin genotypes on key milk production traits

Based on the findings here, sequence variation in exon 3 of *LEP* appeared to have no effect on the average daily milk yield and fat content of milk from these NZ HF×J-cross cows. However, previous studies in cattle have reported that *LEP* variants are associated with variation in milk yield. For example, Clempson *et al.* (2011), investigated the effects of c.239C/T (p.Ala80Val - also identified in this study) with milk yield in 509 Holstein-Friesian heifers in the United Kingdom (UK). Cows with the TT genotype (homozygous for valine at amino acid 80 and equivalent to cows homozygous for variant B_3 here) produced less milk in both their first lactation (milk produced per day) and their second lactation (305 day milk yield), while those with the CC genotype had a lower 305 day milk yield. They suggested that the heterozygote (CT) is the preferred genotype for increasing milk yield. It is however impossible to maintain a viable heterozygotes population without using the two homozygotes, both of which were associated with decreased milk yield.

Even though this study identified the same nucleotide sequence variation as Clempson *et al.* (2011), there was no association between B_3 and milk yield. A possible explanation for this difference may be related to breed difference and/or differences in management practices. While the cows investigated here were crosses of Holstein-Friesian and Jersey cows, albeit of no fixed breed proportion, and raised in a predominantly pasture-based dairy production system, the cows described in the report of Clempson *et al.* (2011) were Holstein-Friesian heifers, raised on farms with a range of management practices representative of those commonly encountered on dairy farms in the UK. Some of these practices include; weaning calves at approximately 4 to 12 weeks of age, and feeding with diets of concentrate (crude protein range 16% to 18%) and forage (such as barley, straw, grass hay, grass silage and maize silage), before they are turned out to pasture at a mean age of 261 days (range of 100 to 431 days) (Brickell *et al.*, 2009). Other studies suggest factors such as breed, lactation-stage, diet, animal body-condition score and the farming environment may affect the production of milk fat (Stelwagen, 2011), while the effects of dairy production system, feeding regime, herd, cow parity and stage of lactation have been suggested to affect milk FA composition (Mele *et al.*, 2016). In this context, it seems plausible that differences in breed and management practices may have been part of the reason for the difference in the findings.

Another reason for the discrepancy between these results and those of Clempson *et al.* (2011) may have been the effect of the number of lactations the cows had experienced, and the main dietary energy supply over the lactations. While the cows investigated here were between 3 to 13 years of age (i.e. not their first time of calving), and the daily milk yield (litres/day) used in this analyses were collected twice daily for a period of 6 months (which likely covered early to mid-lactation stages), the Holstein-Friesian heifers investigated by Clempson *et al.* (2011) were recruited into the study at birth and monitored through their first lactation until the end of their second lactation (or until their death or culling). These heifers calved for the first time at 792 ± 5.3 days, and they may therefore have been younger on average. The milk production traits measured included; average daily milk yield (kg/day), 305-day milk yield (kg) and total milk produced (kg). These values were obtained from monthly records provided by a commercial milk recording services in the UK. Other studies have shown the effect of age at first calving (AFC) on milk yield. Pirlo *et al.* (2000) investigated the 305-day yield records of 1,048,942 Italian Holstein-Friesian heifers and found that the effects of age on milk yield were greater for heifers. Their report revealed that from 36 to 29 months of age at first calving, 170.2 kg of milk were lost, while yield decreased by 254.9 kg from 29 to 24 months at first calving, and 589.8 kg of milk was lost when first calving was at 20 months of age. Although there are no AFC data on the New Zealand cows investigated here to compare with those of Clempson *et al.* (2011), it is possible that the age difference between cows in this study and theirs, may explain the discrepancy in milk yield, especially since the older New Zealand cows have undergone calving a number of times and have also been culled for various reasons, whereas the heifers investigated by Clempson *et al.* (2011) only calved for the first time. Additionally, Serjsen (2005) suggests that milk yield is related to the development of the mammary glands, most of which happens before first calving, thus a lower AFC may be associated with insufficient development of the mammary gland.

Other studies have investigated different regions of bovine *LEP* in different cattle breeds and have reported associations with milk production traits. For example, variation in the promoter region has been associated with milk fat levels in Holstein-Friesian cows (De Matteis *et al.*, 2012), while variations in the intron 2 region have been associated with milk yield in Iranian Holstein and Slovak Simmental cattle (Moussavi *et al.*, 2006; Trakovicka *et al.*, 2015).

Furthermore, the findings here suggest that *LEP* genotype is associated with protein percentage, and that cows carrying the heterozygous genotype A_3B_3 on average have decreased protein levels in their milk. Previous studies have investigated other regions of leptin gene and reported associations with milk protein content. For example, Giblin *et al.* (2010), revealed an effect of the nucleotide variation *LEP*-2470 (c.-2470) in the promoter region of bovine *LEP* on milk protein content in Holstein-Friesian dairy cattle, with the T allele of c.-2470 being associated with reduced milk protein concentration. In another study that investigated the effects of c.73C/T (p.Arg25Cys) in exon 2 with milk protein yield in Holstein cows, the T allele (variant allele for cysteine at amino acid 25) was associated with increased milk protein yield (Buchanan *et al.*, 2003). Chebel *et al.* (2008) also investigated c.73C/T

(p.Arg25Cys) in exon 2, and reported that Holstein cows with the heterozygous CT genotype (i.e. heterozygous for cysteine and arginine at amino acid position 25), had a greater yield of milk protein compared to CC cows (homozygous for arginine at amino acid 25). In other reports, no associations were found between the c.73C/T (p.Arg25Cys) variation and milk production traits in either Polish Black and White or Holstein cattle (Banos *et al.*, 2008; Madeja *et al.*, 2004).

Pirlo *et al.* (2000) reported that age at first calving had a slightly negative effect on first lactation protein percentage in Italian Holstein heifers. Their results suggest that from 36 to 29 months of age at first calving, protein percentage did not increase, but from 29 to 24 months of age, there was an improvement of 0.01%. This was also observed at 24 to 20 months of age. Therefore, it may be that cows with lower AFC tend to have increased milk protein percentage. A possible explanation may lie in the different abilities of young and old heifers to ingest gross forages or concentrates, which can have divergent effects on milk composition (Emery, 1978; Sutton, 1989). Further investigation would therefore appear to be needed to better understand the different effects of age at first calving on protein percentages.

Roche *et al.* (2006) also suggest that in the New Zealand pasture-based dairy system, milk fat percentage increases as the energy supply for milk fat production changes from being sourced from the mobilisation of body fat (in early lactation), to being sourced from feed intake and the *de novo* synthesis of FA (in mid and late lactation), a finding that has also been described by Bionaz *et al.* (2008), who reported that the composition of milk fat changes, as the ratio of *de novo* to 'imported' fat deviates. In this context, as energy supply switches from the mobilisation of body fat to imported energy, the role of leptin and its associated effect on appetite may play a critical role on milk trait performance. This is conceivable, since *LEP* variants have been found to also affect milk FA composition (Pegolo *et al.*, 2016).

4.3.4 Effects of leptin genotype on composition of milk FA

The composition of milk FAs was affected by genotype. Cows carrying the A_3B_3 genotype were observed to have decreased SFA levels, and since leptin has a strong influence on FA metabolism (Sari *et al.*, 2013), the findings in this chapter are in agreement with the report of Pegolo *et al.* (2016). In their investigation, Pegolo *et al.* (2016) examined two SNPs in the intron and upstream region of bovine *LEP* gene respectively in Italian Brown Swiss dairy cows, and reported significant reductions in SFA levels and increases in MUFA and PUFA levels.

The results presented in this chapter are also in line with the findings of Avondo *et al.* (2019), whose report examined the effect of variation at the *LEP* gene intron 1 microsatellite region with milk traits and composition of milk FAs in 16 Girgentana goats. They concluded that the homozygous goats (L genotype) showed lower levels of SFA and higher levels of MUFA and PUFA.

These previous reports, alongside the findings here, are contrary to the reports of Marchitelli *et al.* (2013), whose study did not reveal any significant association between the p.Arg25Cys SNP in *LEP* gene exon 2 and FA traits in Jersey, Piedmontese and Valdostana cattle breeds. Although the results in this chapter did not reveal any significant increase in PUFA level, the observed reduction of SFA in cows carrying the variant genotype (A_3B_3) seems to suggest that this variant genotype may be associated with increased utilization of body fat reserves. This assumption is conceivable, especially considering the fact that a decrease in SFA and an increase in MUFA and PUFA are in concordance with increased mobilization of FAs from adipose tissue (Palmquist *et al.*, 1993; Vrankovic *et al.*, 2017). This view was also shared by Avondo *et al.* (2019), and it's in line with the hypothesis of increased demand for energy as reported by Di Gregorio *et al.* (2014) in their comparative study of goats, sheep, cattle, and water buffalos and the effects of leptin gene Intron 1 microsatellite in goats.

4.3.5 Effect of leptin variants on average body weight

Variant B_3 carrying the c.239C/T (p.Ala80Val) was associated with decreased body weight, although this effect was lost when other variants were included in the model, and when the genotype model was analysed. This is contrary to a previous report by Yang *et al.* (2007), whose study investigated the effects of polymorphisms in exon 3 of bovine *LEP* gene with body weight and body size indexes in Chinese indigenous cattle breeds. Although Yang *et al.* (2007) investigated a 330 bp fragment of bovine *LEP* exon 3, their report did not identify any of the nucleotide sequence variations detected in this study. Their results suggest that cows carrying the variant homozygous genotype *BB,* had increased body weight, body length, heart length, hucklebone width, body height, and average daily gain in Nanyang breed, compared to cows carrying the *AA* or *AB* genotypes. This discrepancy in findings could have arisen from the differences in breed and/or the distribution of genotypes within the sampled population. While the previous report analysed Chinese indigenous cattle breeds carrying the variant homozygous genotypes *BB*, in this study the variant homozygous genotypes B_3B_3 and C_3C_3 in NZ HF × J-cross cows were less than 5% of the population investigated, hence they were not included in the association model. In order to fully appreciate the effect of *LEP* variants on body weight of NZ HF × J-cross cows, further investigations involving larger sample size with possibly higher frequency of homozygous variants, will have to be carried out.

Furthermore, it is important to recognise the correlation between the different lactation stages and , as a consequence of the cow's body fat being used up for this purpose, other biological pathways are affected. For example, this may bring about changes in milk fat composition as well as less energy being partitioned for building up muscle, which in turn, may affect body weight.

4.4 Conclusion

The detection of new as well as previously documented variations in bovine leptin gene indicates a possibility of identifying potential gene marker(s) for the selection of specific traits for increased efficiency in animal production. Although some variants were common across the *Bos taurus* and *Bos indicus* species, which supports the theory of a possible descent from a common ancestor, the physical, physiological and behavioural differences between these two species may have arisen from the different selection pressures exercised over long periods of time by the very different physical environments to which they were exposed.

The findings in this chapter also suggest that cows carrying the variant leptin genotype A_3B_3 are associated with decreased SFA levels in milk. Since heterozygous cows A_3B_3 showed reduced SFA levels, cows with the B_3B_3 genotype might therefore reveal a much lower levels of SFA in milk. Unfortunately, since there were insufficient cattle with the homozygous genotypes B_3B_3 in the samples investigated, further studies involving larger sample sizes across different farms and breeds is needed to validate this claim. However, if the breeding focus is to select for cows with increased weight, caution needs to be taken in selecting cows with either the B_3 variant or the B_3B_3 genotype. This is because, the B_3 variant was observed to cause a decrease in average weight. Although this effect was lost in the presence of other variants, there is still a possibility that cows with the B_3B_3 genotype may have reduced body weight. Additionally, the findings here suggest that leptin variants in exon 3 had no effects on average milk yield and fat, but decreased the protein contents in NZ HF × J-cross cows. Taken together, the findings suggest more research is needed to ascertain whether the effect of *LEP* is time and feed-dependent in lactating cows, and/or whether other genes also affect leptin activity.

Chapter 5

General summary

The main focus of this thesis was to investigate genetic variation in an extended region of the myostatin gene (*MSTN*) and leptin gene (*LEP*) in a variety of cattle breeds from New Zealand and Nigeria, and to ascertain whether these variations influence milk production traits and the composition of milk FAs in NZ HF × J-cross dairy cows.

This is the first research of this kind undertaken in New Zealand, partly because it investigates both New Zealand beef and dairy cattle, as well as the Nigerian dual-purpose cattle breeds. Also, it is the first research exploring associations of *MSTN* variants with the composition of milk FAs and milk production traits in NZ HF × J-cross cows.

In summary, a total of eight regions (Five on *MSTN* and three on *LEP*) were investigated in fourteen different breeds of beef, dairy and dual-purpose cattle from New Zealand and Nigeria. These included the NZ Hereford, Angus, South Devon, Composite, Charolais, Red Poll, Shorthorn, Simmental, Murray Grey, NZ HF x J-cross cows, as well as the Nigerian dual purpose Sokoto Gudali, Red Bororo, White Fulani and cross-bred Holstein Friesian × White Fulani cattle. In the *MSTN,* five pairs of primers were used to genotype 883 cattle from fourteen breeds. A total of five amplicons covering 1,967 bp were generated. The amplified regions included exons 1, 2 and 3 and parts of introns 1 and 2. Eighteen single-nucleotide variations were detected and twenty haplotypes (H1 to H20) were resolved across *MSTN* introns 1 and 2 regions. H1 to H8 were found in all the New Zealand and Nigerian breeds, whereas H19 and H20 were only found in the Nigerian breeds.

For the *LEP*, three pairs of primers were used to genotype 657 cattle from eight different breeds from New Zealand and Nigeria. Three amplicons which included exons 2 and 3, and parts of introns 1 and 2 covering 1,152 bp of non-contiguous sequence were generated. Twelve nucleotide variations were identified across these regions examined. The haplotypes for *LEP* could not be resolved due to the small sample size and the unavailability of enough homozygous samples to enable determination of haplotypes.

In the *MSTN* regions investigated, three variants (*A, B, C*) in intron 1 and three variants (A_4, B_4, D_4) in intron 2 were common in both the New Zealand *Bos taurus* and the Nigerian *Bos indicus* breeds investigated. Similarly, in the leptin gene examined, variant B_2 in the intron 2 region was observed in all the New Zealand and Nigerian cattle breeds investigated. This suggests that these variations are shared across cattle of *Bos taurus* and *Bos indicus* (African Zebu cattle) origin, which likely reflects the close relationship between these two species in terms of origin. Literature has suggested that these two species may have originated from a common ancestor, and that the auroch (*Bos primigenius*), is the progenitor of all taurine and zebu (African *Bos indicus*) cattle (Edwards *et al.*, 2007). The

estimates of divergence times for *Bos taurus* and *Bos indicus* from a common ancestor are all pre-Neolithic (Hiendleder *et al.*, 2008). The sharing of common haplotypes suggest that cattle studied here might be more closely related, although it does not provide definitive proof. The suggestion is supported by Dunner *et al.* (2003), who suggested that the pattern of haplotype sharing is an indicator of the history of the different bovine breeds, and thus the distribution of shared haplotypes is useful in describing population relationships.

The eleven novel nucleotide sequence variations in this thesis were all identified in the intronic regions of *MSTN* and *LEP* genes. Recently, introns have become far more important to our understanding of gene and genome functionality, since they are able to encode regulatory elements (Yutzey *et al.*, 1989) that participate in splicing, transcription, and recombination activities. In this context, while there is no report on the possible effect of the absence of bovine *MSTN* intron 1 on gene expression, it would be worthwhile exploring this path, especially, since previous study (McPherron *et al.*, 1997) have only focused on examining the effects of 'knocking-out' the functional portion of the MSTN peptide. Also, since nucleotide sequence variation at close proximity to the intron/exon boundaries may affect gene expression (Sjakste *et al.*, 2011), validating this in the context of this thesis will require future studies on MSTN expression in cattle carrying the observed c.748-78delT deletion and the same for leptin expression in cattle carrying the *LEP* c.144+42delC deletion.

The results for associations between *MSTN* genotypes and milk FAs composition could be interpreted to suggest that cows with the *CD* genotype might produce the 'preferred' FA profile in milk. A previous report in dual purpose Belgian Blue (DP-BB) cows has suggested that one copy of the so-called 'wild-type + allele' of *MSTN* is responsible for higher milk, protein, and fat yields, whereas a single copy of the mutant '*mh*' allele decreases the SFA content of milk from DP-BB (Buske *et al.*, 2011). However, since MSTN is primarily known for its role in regulating muscle growth and development, a more detailed investigation into its physiology, pathway and mode of action is still required to unravel its pleiotropic effect on milk fat composition.

Interestingly, variation in the exon 3 of *LEP* was associated with a decrease in the SFA content of milk, but had no effect on average milk yield and fat contents in NZ HF × J-cross cows. This will require further investigation into when leptin is expressed, it's interaction with other genes or QTLs, and to what extent diet and the stage of lactation influence the gene's effect on milk components.

Based on the findings in this thesis, variation in *MSTN* and *LEP* may be employed as selection tools for increasing the concentration of PUFA levels and decreasing SFA levels in milk of NZ HF x J-cross cows in New Zealand. Future studies need to be carried out to validate the effect of genotypes on composition of milk FAs in larger sample sizes across different farms, breeds and cross-breeds.

Since improving the efficiency of cattle production systems will entail selecting for fast growing animals with increased muscling, and desirable maternal reproductive traits, the findings presented

here and the information that could arise from future studies will enable a better understanding of the pleotropic effects of genes on muscularity and adiposity, and this could improve the efficiency and selection precision in animal production systems.

References

Aguilera, A., and Gomez-Gonzalez, B. (2008). Genome instability: A mechanistic view of its causes and consequences. *Nature Reviews Genetics*, 9 (3), 204-17.

Ahima, R.S., Saper, C. B., Flier, J. S., and Elmquist, J.K. (2000). Leptin regulation of neuroendocrine systems. *Frontiers in Neuroendocrinology,* 21, 263-307.

Aiello, D., Patel, K., and Lasagna, E., 2018. The myostatin gene: an overview of mechanisms of action and its relevance to livestock animals. *Animal Genetics*, 49, 505-519.

Andersson, L. (2001). Genetic dissection of phenotypic diversity in farm animals. *Nature Reviews Genetics,* 2(2), 130-138.

Amico, J.A., Thomas, A., Crowley, R.S. and Burmeister, L.A. (1998). Concentrations of leptin in the serum of pregnant, lactating, and cycling rats and of leptin messenger ribonucleic acid in rat placental tissue. *Life Sciences,* 63, 1387-1395.

Amstalden, M., Garcia, M.R., Williams, S. W., Stanko, R. L., Nizielski, S. E., Morrison, C. D., Keisler, D. H., and Williams, G.L. (2000). Leptin gene expression, circulating leptin, and luteinizing hormone pulsatility are acutely responsive to short-term fasting in prepubertal heifers: relationships to circulating insulin and insulin-like growth factor I (1). *Biology of Reproduction* 63, 127-133.

Amstalden, M., Zieba, D. A., Edwards, J. F., Harms, P. G., Welsh, Jr. T. H., Stanko, R. L., and Williams, G. L. (2003). Leptin acts at the bovine adenohypophysis to enhance basal and gonadotropin-releasing hormone-mediated release of luteinizing hormone: differential effects are dependent upon nutritional history. *Biology of Reproduction,* 69, 1539-1544.

Arnould, V.M.R and Soyeurt, H. (2009). Genetic variability of milk FAs. *Journal of Applied Genetics,* 209, 29-39.

Avondo, M., Trana, A.D., Valenti, B., Criscione, A., Bordonaro, S., Angelis, A.D., Giorgio, D., and Gregorio, P.D. (2019). Leptin Gene Polymorphism in Goats Fed with Diet at Different Energy Level: Effects on Feed Intake, Milk Traits, Milk FAs Composition, and Metabolic State. *Animals*, 9, 424.

Aznarez, I., Barash, Y., Shai, O., He, D., Zielenski, J., Tsui, L.C., Parkinson, J., Frey, B.J., Rommens, J.M., and Blencowe, B.J. (2008). A systematic analysis of intronic sequences downstream of 5' splice sites reveals a widespread role for U-rich motifs and TIA1/TIAL1 proteins in alternative splicing regulation. *Genome Research,* 18, 1247-1258.

Banos, G., Woolliams, J.A., Woodward, B.W., Forbes, A.B., Coffey, M.P. (2008) Impact of Single Nucleotide Polymorphisms in Leptin, Leptin Receptor, Growth Hormone Receptor, and Diacylglycerol Acyltransferase (DGAT1) Gene Loci on Milk Production, Feed, and Body Energy Traits of UK Dairy Cows. *Journal of Dairy Science,* 91, 3190-3200.

Barb, C. R., Yan, X., Azain, M. J., Kraeling, R. R., Rampacek, G. B., and T. G. Ramsay. (1998). Recombinant porcine leptin reduces feed intake and stimulates growth hormone secretion in swine. *Domestic Animal Endocrinology,* 15, 77-86.

Bauman, D.E., and Bruce Currie, W. (1980). Partitioning of nutrients during pregnancy and lactation: a review of mechanisms involving homeostasis and homeorhesis. *Journal of Dairy Science*, 63(9), 1514-1529.

Bellinge, R.H., Liberles, D.A., Iaschi, S.P., O'Brien P, A., and Tay, G.K. (2005). Myostatin and its implications on animal breeding: a review. *Animal Genetics*, 36, 1-6.

Bionaz, M., and Loor, J. (2008b). Gene networks driving bovine milk fat synthesis during the lactation cycle. *BMC Genomics*, 9(1), 366.

Blache, D., Tellam, R. L., Chagas, L. M., Blackberry, M. A., Vercoe, P. E., and Martin, G. B. (2000). Level of nutrition affects leptin concentrations in plasma and cerebrospinal fluid in sheep. *Journal of Endocrinology,* 165, 625-637.

Block, S.S., Butler, W.R., Ehrhardt, R.A., Bell, A.W., Van Amburgh, M.E., and Boisclair, Y.R. (2001). Decreased concentration of plasma leptin in periparturient dairy cows is caused by negative energy balance. *Journal of Endocrinology*, 171, 339-348.

Blott, S., Williams, J., and Haley, C. (1998). Genetic variation within the Hereford breed of cattle. *Animal Genetics*, 29, 202-211.

Blott, S., Kim, J. J., Moisio, S., Schmidt-Kuntzel, A., Cornet, A., Berzi, P., Cambisano, N., Ford, C., Grisart, B., Johnson, D., Karim, L., Simon, P., Snell, R., Spelman, R., Wong, J., Vilkki, J., Georges, M., Farnir, F., and Coppieters, W. (2003). Molecular dissection of a quantitative trait locus: a phenylalanine-to-tyrosine substitution in the transmembrane domain of the bovine growth hormone receptor is associated with a major effect on milk yield and composition. *Genetics,* 163, 253-266.

Boman, I.A., Klemetsdal, G., Blichfeldt, T., Nafstad, O., and Vage, D.I. (2009). A frameshift mutation in the coding region of the myostatin gene (MSTN) affect carcass conformation and fatness in Norwegian White Sheep (Ovis aries). *Animal Genetics,* 40, 418-422.

Boman, I.A., and Vage, D.I. (2009). An insertion in the coding region of the myostatin (MSTN) gene affects carcass conformation and fatness in the Norwegian Spaelsau (Ovis aries). *BMC Research Notes,* 2, 98.

Boman, I.A., Klemetsdal, G., Nafstad, O., Blichfeldt T., and Våge, D.I. (2010). Selection based on progeny testing induces rapid changes in myostatin allele frequencies – a case study in sheep. *Journal of Animal Breeding and Genetics*, 128, 1, 52-55.

Bonnet, M., I. Gourdou, C. Leroux, Y. Chilliard, and J. Djiane. (2002). Leptin expression in the ovine mammary gland: putative sequential involvement of adipose, epithelial, and myoepithelial cells during pregnancy and lactation. *Journal of Animal Science*, 80, 723-728.

Bradley, D.G., Machugh, D.E., Cunningham, P., and Loftus, R.T. (1996). Mitochondrial diversity and the origins of African and European cattle. *Proceedings of the National Academy of Sciences*, USA 93, 5131-5135.

Brickell, J.S., Bourne, N., McGowan, M.M., and Wathes, D.C. (2009). Effect of growth and development during the rearing period on the subsequent fertility of nulliparous Holstein-Friesian heifers. *Theriogenology* 72, 408-416.

Brogan, R.S., Mitchell, S.E., Trayhurn, P., and Smith, M.S. (1999). Suppression of leptin during lactation: contribution of the suckling stimulus versus milk production. *Endocrinology,* 140, 2621-2627.

Brogan, R. S., Grove, K. L., and Smith, M.S. (2000). Differential regulation of leptin receptor but not orexin in the hypothalamus of the lactating rat. *Journal of Neuroendocrinology,* 12, 1077-1086.

Buchman, A.R., and Berg, P. (1988). Comparison of intron-dependent and intron-independent gene expression. *Molecular and Cell Biology*, 8(10), 4395-4405.

Buchanan, F. C., Fitzsimmons, C. J., van Kessel, A. G., Thue, T. D., Winkelman-Sim, D. C., and Schmutz, S.M. (2002). Association of a missense mutation in the bovine leptin gene with carcass fat content and leptin mRNA levels. *Genetics Selection Evolution,* 34, 105-116.

Buchanan, F.C., Van Kessel, A.G., Waldner, C., Christensen, D.A., Laarveld, B., and Schmutz, S.M. (2003). Hot topic: an association between a leptin single nucleotide polymorphism and milk and protein yield. *Journal of Dairy Science*, 86, 3164-3166.

Buske, B., Gengler, N., and Soyeurt, H. (2011). Short-communication: influence of the muscle hypertrophy mutation of the myostatin gene on milk production traits and milk FA composition in dual-purpose Belgian Blue dairy cattle. *Journal of Dairy Science,* 94, 3687-3692.

Byun, S., Fang, Q., Zhou, H., and Hickford, J. (2009). An effective method for silver-staining DNA in large numbers of polyacrylamide gels. *Analytical Biochemistry*, 385(1), 174-175.

Campfield, L. A., Smith, F. J., Guisez, Y., Devos, R., and Burn, P. (1995). Recombinant mouse OB protein: evidence for a peripheral signal linking adiposity and central neural networks. *Science,* 269, 546-549.

Chasin, L. (2007). Searching for splicing motifs. *Advances in Experimental Medicine and Biology,* 623, 85-106.

Charlier, C., Coppieters, W., Farnir, F., Grobet, L., Leroy, P., Michaux, C., M. Mni, M., Schwers, A., Vanmanshoven, P., Hanset, R., Georges, M., (1995). The mh gene causing double-muscling in cattle maps to bovine chromosome 2. *Mammalian Genome*, 6(11), 788-792.

Chebel, R.C., Susca, F., Santos, J.E. (2008). Leptin genotype is associated with lactation performance and health of Holstein cows. *Journal of Dairy Science*, 91, 2893-2900.

Chilliard, Y., Ferlay, A., Faulconnier, Y., Bonnet, M., Rouel, J., and Bocquier, F., (2000). Adipose tissue metabolism and its role in adaptations to undernutrition in ruminants. *Proceedings of Nutrition Society,* 59, 127-134.

Chilliard, Y., Ferlay, A., and Doreau, M. (2001). Effect of different types of forages, animal fat or marine oils in cow's diet on milk fat secretion and composition, especially conjugated linoleic acid (CLA) and polyunsaturated FAs. *Livestock Production Science*, 70(1), 31-48.

Chilliard, Y., Delavaud, C., and Bonnet, M. (2005). Leptin expression in ruminants: Nutritional and physiological regulations in relation with energy metabolism. *Domestic Animal Endocrinol*ogy, 29, 3-22.

Considine, R.V., Considine, E. L., Williams, C. J., Hyde, T.M., and Caro, J.F. (1996). The hypothalamic leptin receptor in humans: identification of incidental sequence polymorphisms and absence of the db/db mouse and fa/fa rat mutations. *Diabetes,* 45, 992-994.

Cho, T.J., Gerstenfeld, L.C., Einhorn, T.A. (2002). Differential temporal expression of members of the transforming growth factor beta superfamily during murine fracture healing. *Journal of Bone and Mineral Research,* 17, 513-520.

Chung, S., and Perry, R. (1989). Importance of introns for expression of mouse ribosomal protein gene rpL32. *Molecular and Cell Biology,* 9(5), 2075-2082.

Chris, S., James Z., and Chris, N. (2010). Dairy's role in sustaining New Zealand - the sector's contribution to the economy. New Zealand institute of economic research. https://www.academia.edu/20524799/Dairy_s_Role_in_Sustaining_New_Zealand_the_sector_s_contribution_to_the_economy (accessed online 2oth March, 2018).

Clempson, A.M., Pollott, G.E., Brickell, J.S., Bourne, N.E., Munce, N., and Wathes, D.C. (2011). Evidence that leptin genotype is associated with fertility, growth, and milk production in Holstein cows. *Journal of Dairy Science*, 94, 3618-3628.

Clop, A., Marcq, F., Takeda, H., Pirottin, D., Tordoir, X., Bibé, B., Bouix, J., Caiment, F., Elsen, J.M., Eychenne, F., Larzul, C., Laville, E., Meish, F., Milenkovic, D., Tobin, J., Charlier, C., and Georges, M. (2006). A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nature Genetics*, 38(7), 813-818.

DairyNZ Quickstats. (2017). New Zealand Dairy Statistics 2016-2017. Livestock Improvement Corporation Ltd & DairyNZ Ltd. https://www.dairynz.co.nz/publications/dairy-industry/new-zealand-dairy-statistics-2016-17/ (accessed online 3rd February, 2019).

DairyNZ Quickstats. (2018a). DairyNZ Economic Survey 2017-2018. DairyNZ Ltd. https://www.dairynz.co.nz/publications/dairy-industry/dairynz-economic-survey-2017-18/ (accessed online 17th October, 2019)

DairyNZ Quickstats. (2018b). New Zealand Dairy Statistics 2017-2018. Livestock Improvement Corporation Ltd & DairyNZ Ltd. https://www.dairynz.co.nz/media/5791052/quickstats-about-dairying-new-zealand-2019.pdf (accessed online 11th November, 2019).

DairyNZ Quickstats. (2019). Dairy sector quick facts: Farms, economics, Markets. https://www.dairynznewslink.co.nz/wp-content/uploads/2019/06/Dairy-sector-quick-facts-June-2019-2.pdf (accessed 8th January, 2020).

Delavaud, C., Bocquier, F., Chilliard, Y., Keisler, D.H., Gertler, A., and Kann, G. (2000). Plasma leptin determination in ruminants: effect of nutritional status and body fatness on plasma leptin concentration assessed by a specific RIA in sheep. *Journal of Endocrinology,* 165, 519-526.

Delavaud, C., Ferlay, A., Falconnier, Y., Bocquier, F., Kann, G., and Chilliard, Y. (2002). Plasma leptin concentration in adult cattle: Effect of breed, adiposity, feeding level, and meal intake. *Journal of Animal Science,* 80, 1317-1328.

De Matteis, G., Scatà, M.C., Grandoni, F., Petrera, F., Abeni, F., Catillo, G., Napolitano, F., and Moioli, B. (2012). Association analyses of single nucleotide polymorphisms in the leptin and leptin

receptor genes on milk and morphological traits in Holstein cows. *Open Journal of Animal Sci*ence, 2, 174-182.

Destremau, K., and Siddharth, P. (2018). How does the dairy sector share its growth?: An analysis of the flow-on benefits of dairy's revenue generation. NZIER report to dairy companies association of New Zealand. (https://nzier.org.nz/static/media/filer_public/9f/0e/9f0e40ea-0178-4ef9-950f-5546ef483eec/dcanz_2018_final.pdf) (Accessed online 12th December, 2019).

Dewhurst, R. J., Shingfield, K.J., Lee, M.R.F., and Scollan, N.D. (2006). Increasing the concentrations of beneficial polyunsaturated FAs in milk produced by dairy cows in high-forage systems. *Animal Feed Science and Technology*, 131(3), 168-206.

Di Gregorio, P., Di Trana, A., Celi, P., Claps, S., and Rando, A. (2014) Comparison of goat, sheep, cattle and water buffalo leptin (*LEP*) genes and effects of the Intron 1 microsatellite polymorphism in goats. *Animal Production Science*, 54, 1258-1262.

Du, R., Chen, Y.F., An, X.R., Yang, X.Y., Ma, Y., Zhang, L., Yuan, X.L., Chen, L.M., and Qin, J. (2005). Cloning and sequence analysis of myostatin promoter in sheep. *DNA Sequence*, 6, 412-417.

Dunner, S., Charlier, C., Farnir, F., Brouwers, B., Canon, J., and Georges, M. (1997). Towards interbreed IBD fine mapping of the mh locus: double-muscling in the Asturiana de los Valles breed involves the same locus as in the Belgian Blue cattle breed. *Mammalian genome*, 8(6), 430-435.

Dunner, S., Miranda, M.E., Amigues, Y., Cañón, J., Georges, M., Hanset, R., Williams, J. L., and Ménissier, F. (2003). Haplotype diversity of the myostatin gene among beef cattle breeds. *Genetics Selection and Evolution,* 35, 103-118.

Dunner, S., Sevane, N., García, D., Cortés, O., Valentini, A., Williams, J. L., Mangin, B., Canon, J., Levéziel, H. (2013). Association of genes involved in carcass and meat quality traits in 15 European bovine breeds. *Livestock Science*, 154(13), 34-44.

Dyer, C.J., Simmons, J.M., Matteri, R.L., and. Keisler, D.H. (1997). Leptin receptor mRNA is expressed in ewe anterior pituitary and adipose tissues and is differentially expressed in hypothalamic regions of well- fed and feed-restricted ewes. *Domestic Animal Endocrinology,* 14, 119-128.

Edwards, C. J., Bollongino, R., Scheu, A., Chamberlain, A., Tresset, A., Vigne, J.-D., Baird, J.F.G., Larson, S.Y., Ho, S.Y.W., Heupink, T. H., Shapiro, B., Freeman A.R., Thomas M.G., Arbogast, R., Arndt B., Bartosiewicz L., Benecke N, Budja M., Chaix, L., Choyke, A.M., Coqueugniot, E., Döhle H., Göldner H, Hartz S., Helmer D., Herzig B., Hongo H., Mashkour, M., Özdogan M., Pucher, E., Roth, G., Schade-Lindig, S., Schmölcke U., Schulting, R.J, Stephan, E., Uerpmann H, Vörös, I., Voytek, B., Bradley D.G., and Burger J. (2007) . Mitochondrial DNA analysis shows a Near Eastern

Neolithic origin for domestic cattle and no indication of domestication of European aurochs. *Proceedings of the Royal Society of London B: Biological Sciences,* 274, 1377-1385.

Ehrhardt, R. A., Slepetis, R. M., Siegal-Willott, J., Van Amburgh, M. E., Bell, A.W. and Boisclair, Y. R. (2000). Development of a specific radioimmunoassay to measure physiological changes of circulating leptin in cattle and sheep. *Journal of Endocrinology*, 166, 519-528.

Ehrhardt, R. A., Slepetis, R.M., Bell, A.W., and Boisclair, Y.R. (2001). Maternal leptin is elevated during pregnancy in sheep. *Domestic Animal Endocrinology,* 21, 85-96.

Epstein, H., (1973). The Origin of the Domestic Animals of Africa. *Journal of Mammalogy*, 54(1), 305-306, https://doi.org/10.2307/1378907.

Epstein, H., and Mason, I.L. (1984). Evolution of Domesticated Animals, Longman, London Cattle, 627.

Esmailizadeh, A. K., Bottema, C.D.K., Sellick, G.S., Verbyla, A.P., Morris, C.A., Cullen, N.G., and Pitchford, W.S. (2008). Effects of the myostatin F94L substitution on beef traits. *Journal of Animal Science*, 86, 1038-1046.

Ensembl (EMBL-EBI, Hinxton, United Kingdom (UK), and Release 93 - July 2018) UMD3.1:GK000002.2 for cow MSTN (https://asia.ensembl.org/index.html).

Ensembl (EMBL-EBI, Hinxton, United Kingdom, and release 96 - April 2019) ARS-UCD1.2 for cow Leptin gene (https://www.ncbi.nlm.nih.gov/gene/280836).

Emery, R.S. (1978). Feeding for increased milk protein. *Journal of Dairy Science*, 61, 825-828.

Ellis, L.A., Mastro, A.M., and Picciano, M.F. (1997) Do milk-borne cytokines and hormones influence neonatal immune cell function? *The Journal of Nutrition*, 127, 985S–988S.

Feuermann, Y., Mabjeesh, S.J., Shamay, A. (2004). Leptin affects prolactin action on milk protein and fat synthesis in the bovine mammary gland. *Journal of Dairy Science,* 87, 2941-2946.

Fitzsimmons, C. J., Schmutz, S. M., Bergen, R.D., and McKinnon, J.J. (1998). A potential association between the BM 1500 microsatellite and fat deposition in beef cattle. *Mammalian Genome,* 9, 432-434.

Forhead, A. J., Thomas, L., Crabtree, J., Hoggard, N., Gardner, D.S., Giussani, D.A., and Fowden, A.L., (2002). Plasma leptin concentration in fetal sheep during late gestation: ontogeny and effect of glucocorticoids. *Endocrinology,* 143, 1166-1173.

Gan, S.Q., Du, Z., Liu, S. R., Yang, Y. L., Shen, M., Wan, X. H., Yin, J. L., Hu, X. X., Feifan, J.J., Wang, J. H., He, Q. H., Zhang, Y. and Sand, Li. N. (2008). Association of SNP haplotypes at the 66 myostatin gene with muscular hypertrophy in sheep. *Asian-Australian Journal of Animal Science*, 21 (7), 928-935.

Garnsworthy, P. C., Masson, L. L., Lock, A. L., and Mottram, T. T. (2006). Variation of milk citrate with stage of lactation and De Novo FA synthesis in dairy cows. *Journal of Dairy Science*, 89(5), 1604-1612.

Garcia, M. D., Casanueva, F.F., Dieguez, C., and Senaris, R.M. (2000). Gestational profile of leptin messenger ribonucleic acid (mRNA) content in the placenta and adipose tissue in the rat, and regulation of the mRNA levels of the leptin receptor subtypes in the hypothalamus during pregnancy and lactation. *Biology of Reproduction,* 62, 698-703.

Georges, M., Nielsen, D., Mackinnon, M., Mishra, A., Okimoto, R., Pasquino, A.T., Sargeant, L.S., Sorensen, A., Steele, M.R., Zhao, X., Womack, J.E., Hoeschele, I. (1995). Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. *Genetics,* 139, 907-920.

Giblin, L., Butler, S.T., Kearney, B.M., Waters, S.M., Callanan, M.J., Berry, D.P. (2010) Association of bovine leptin polymorphisms with energy output and energy storage traits in progeny tested Holstein-Friesian dairy cattle sires. *BMC Genetics*, 11, 73.

Gong, H., Zhou, H., Hickford, J.G.H. (2011). Diversity of the glycine/tyrosine-rich keratin-associated protein 6 gene (KAP6) family in sheep. *Molecular Biology and Reproduction,* 38(1), 31-35.

González-Cadavid N., Taylor, W., Yarasheski, K., Sinha-hikim I., Ma K., Ezzat, S., Shen, R., Lalani, R., Asa, S., Mamita, M., Nair, G., Arver, S., Bhasin, S. (1998). Organization of the human myostatin gene and expression in healthy men and HIV-infected men with muscle wasting. *Proceedings of National Academy of Science*, USA; 95 14938-14943.

Grisolia, A.B., Angelo, G.T.D., Porto Neto, L.R., Siqueira, F., and Garcia, J.F., (2009). Myostatin (GDF8) single nucleotide polymorphisms in Nellore cattle. *Genetics and Molecular Research* 8, 822-830.

Grobet L., Royo Martin L.J., Poncelet D., Pirottin D., Brouwers B., Riquet J., Schoeberlein A., Dunner S., Menissier F., Massabanda J., Fries R., Hanset R., Georges M. A. (1997). Deletion in the myostatin gene causes double muscled phenotype in cattle. *Nature Genetics*, 17, 71-74.

Grobet, L., Poncelet, D., Royo, L.J., Brouwers, B., Pirottin, D., Michaux, C., Me´nissier, F., Zanotti, M., Dunner, S., and Georges, M. (1998). Molecular definition of an allelic series of mutations

disrupting the myostatin function and causing double-muscling in cattle. *Mammalian Genome,* 9, 210-213.

Grisart, B., Coppieters, W., Farnir, F., Karim, L., Ford, C., Berzi, P., Cambisano, N., Mni, M., Reid, S., Simon, P., Spelman, R., Georges, M., Snell, R. (2002). Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. *Genome Research*, 12(2), 222-231.

Grummer, R.R. (1991). Effects of feed on the composition of milk fat. *Journal of Dairy Science,* 74, 3244-3257.

Haegeman, A., Van Zeveren, A., and Peelman, L. J. (2000). New mutation in exon 2 of the bovine leptin gene. *Animal Genetics,* 31, 79.

Hager, J., Clement, K., Francke, S., Dina, C., Raison, J., Lahlou, N., Rich, V., Pelloux, V., Basdevant, A., Guy-Grand, B., North, M., and Froguel, P. (1998). A polymorphism in the 5' untranslated region of the human ob gene is associated with low leptin levels. *International Journal of Obesity and Related Metabolic Disorder*, 22, 200-205.

Hajihosseinlo, A., Hashemi, A., Sadeghi, S. (2012). Association between polymorphism in exon 3 of leptin gene and growth traits in the Makooei sheep of Iran. *Livestock Research for Rural Development*, 24, 543-546.

Halaas, J. L., Gajiwala, K.S., Maffei, M., Cohen, S.L., Chait, B.T., Rabinowitz, D., Lallone, R. L., Burley, S.K., and Friedman, J.M. (1995). Weight-reducing effects of the plasma protein encoded by the obese gene. *Science,* 269, 543-546.

Han, J., Forrest, R. H., Sedcole, J. R., and Hickford, J.G.H. (2015). Myostatin (MSTN) gene haplotypes and their association with growth and carcass traits in New Zealand Romney lambs. *Small Ruminant Research*, 127, 8-19.

Harris, B.L., Clark, J.M., and Jackson, R.G. (1996). Across breed evaluation of dairy cattle. *Proceedings of the New Zealand Society of Animal Production*, 56. http://www.nzsap.org/system/files/proceedings/1996/ab96005.pdf (accessed 8th August, 2018).

Haruna I.L., Ekegbu, U.J., Ullah, F., Amirpour-Najafabadi, H., Zhou, H., and Hickford, J.G.H. (2020). Genetic variation and haplotypic diversity in the myostatin gene of New Zealand cattle breeds. *Gene*, 740, available online 25th May, 2020.

Hayashi, K. (1991). PCR-SSCP: A simple and sensitive method for detection of mutations in the genomic DNA. *Genome Research*, 1, 34-38.

He, Y.L., Wu, Y.H., Quan, F.S. Liu, Y.G., and Zhang, Y. (2010). Molecular analysis of the first intron in the bovine myostatin gene. *Molecular Biology Reports,* 38, 4643-4649.

He, Y.L., Wu, Y.H., Quan, F.S. Liu, Y.G., and Zhang, Y. (2013). Comparative analysis of myostatin gene and promoter sequences of Qinchuan and Red Angus cattle. *Genetics and Molecular Research*, 12 (3), 3398-3406.

Heaton, M.P., Smith, T.P.L., Freking, B.A., Wray, J.E., and Leymaster, K.A. (2007). Characterization of SNPs in the myostatin 3'UTR region of U. S. sheep (unpublished paper). NCBI number EF182738

Hennighausen, L., and Robinson, G.W. (2001). Signalling pathways in mammary gland development. *Developmental Cell*, 1, 467-475.

Henry, B. A., Goding, J. W., Alexander, W. S., Tilbrook, A. J., Canny, B. J., Dunshea, F., Rao, A., Mansell, A., and Clarke, I. J. (1999). Central administration of leptin to ovariectomized ewes inhibits food intake without affecting the secretion of hormones from the pituitary gland: evidence for a dissociation of effects on appetite and neuroendocrine function. *Endocrinology,* 140, 1175-1182.

Henry, B. A., Goding, J.W., Tilbrook, A.J., Dunshea, F.R., and Clarke, I.J. (2001). Intra-cerebroventricular infusion of leptin elevates the secretion of luteinising hormone without affecting food intake in long-term food- restricted sheep, but increases growth hormone irrespective of bodyweight. *Journal of Endocrinology*, 168, 67-77.

Hey, J. (2004). What's So Hot about Recombination Hotspots? PLoS Biol. 2(6): e190.

Heyen, D.W., Weller, J.I. Ron, M., Band, M., Beever, J.E., Feldmesser, E., Da, Y., Wiggans, G.R., Van Raden, P.M. and Lewin, H.A. (1999). A genome scan for QTL influencing milk production and health traits in dairy cattle. *Physiological Genomics*, 1, 165-175.

Hickford, J. G. H., Forrest, R. H., Zhou, H., Fang, Q., Han, J., Frampton, C. M., and Horrell, A.L. (2010). Polymorphisms in the ovine myostatin gene (MSTN) and their association with growth and carcass traits in New Zealand Romney sheep. *Animal Genetics*, 41, 64-72.

Hiendleder, S., Lewalski, H., and Janke, A. (2008). Complete mitochondrial genomes of *Bos taurus* and *Bos indicus* provide new insights into intra-species variation, taxonomy and domestication. *Cytogenetics Genome Research*, 120, 150-156.

Hiller, M., Zhang, Z., Backkofen, R., and Stamm, S. (2007). PremRNA Secondary structures influence exon recognition. *PLoS Genetics,* 3, e204.

Hiroike, T., Higo, J., Jingami, H., and Toh, H. (2000). Homology modeling of human leptin/leptin receptor complex. *Biochemical and Biophysical Research Communications*, 275, 154-158.

Houseknecht, K. L., Baile, C.A., Matteri, R.L., and Spurlock, M.E. (1998). The biology of leptin: a review. *Journal of Animal Science*, 76, 1405-1420.

Hunt, R.C., Simhadri, V.L., Iandoli, M., Sauna, Z.E., and Kimchi-Sarfaty, C. (2014). Exposing synonymous mutations. *Trends in Genetics*, 30, 308-321.

Influx holding (2017). Feeding high yield milk cows while maintaining milk yield and fertility https://www.influxlipids.com/single-post/2017/08/12/Feeding-High-Yield-Milk-Cows-While-Maintaining-Milk-Yield-And-Fertility (accessed 24th April, 2019).

James, W., and John, H. (1999). New Zealand pasture and crop science: Auckland, N.Z.; Oxford: Oxford University Press.

Jeanplong, F., Sharma, M., Somers, W.G., Bass, J.J., and Kambadur, R. (2001) Genomic organization and neonatal expression of the bovine myostatin gene. *Molecular and Cellular Biochemistry*, 220(1-2), 31-37.

Jeffreys, A.J., Neumann, R., Panayi, M., Myers, S., Donnelly, P. (2005). Human recombination hot spots hidden in regions of strong marker association. *Nature Genetics*, 37, 601-606.

Ji, S., Losinsky, R.L., Cornelius, S.G., Frank, G.R., Willis, G.M., Gerrard, D.E., Depreux, F.F., and Spurlock, M.E. (1998). Myostatin expression in porcine tissues: tissue specificity and developmental and postnatal regulation. *American Journal of Physiology – Regulatory, Integrative and Comparative Physiology*, 275, 1265-1273.

Jiang, Z.H., and Gibson, J.P. (1999). Genetic polymorphisms in the leptin gene and their association with fatness in four pig breeds. *Mammalian Genome*, 10, 191-193.

Jin, L., Zhang, S., Burguera, B.G., Couce, M.E., Osamura, R.Y., Kulig, E., and Lloyd, R.V. (2000). Leptin and leptin receptor expression in rat and mouse pituitary cells. *Endocrinology,* 141, 333-339.

John, B., and Pambudi, D. (2017). Dairy trade's economic contribution to New Zealand. https://nzier.org.nz/static/media/filer_public/29/33/29336237-3350-40ce-9933-a5a59d25bd31/dairy_economic_contribution_update_final_21_february_2017.pdf (accessed 17th January, 2019).

Jonas, E., Martin, G.B., Celi, P., Soattin, M., Thomson, P.C., and Raadsma, H.W (2016) Association of polymorphisms in leptin and leptin receptor genes with circulating leptin concentrations, production and effciency traits in sheep. *Small Ruminant Research*, 136, 78-86.

Kadokawa, H., Blache, D., Yamada, Y., and Martin, G.B. (2000). Relationships between changes in plasma concentrations of leptin before and after parturition and the timing of first postpartum ovulation in high-producing Holstein dairy cows. *Reproduction Fertility and Development*, 12, 405-411.

Kambadur, R., Sharma, M., Smith, T. P., and Bass, J. J. (1997). Mutations in myostatin (GDF8) in double-muscled Belgian Blue and Piedmontese cattle. *Genome Research*, 7, 910-916.

Kantanen, J., Olsaker, I., Holm, L.-E., Lien, S., Vilkki, J., Brusgaard, K., Eythorsdottir, E., Danell, B., and Adalsteinsson, S. (2000b). Genectic Diversity and Population Structure of 20 North European Cattle Breeds. *Journal of Heredity*, 91, 446-457.

Kawaguchi, F., Okura, K., Oyama, K., Mannen, H., and Sasazaki, S. (2017). Identification of leptin gene polymorphisms associated with carcass traits and FA composition in Japanese Black cattle. *Animal Scienece Journal*, 88, 433-438.

Kijas, J.M., McCulloch, R., Hocking Edwards, J.E., Oddy V.H., Lee, S.H. and Van der Werf, J. (2007). Evidence for multiple alleles effecting muscling and fatness at the ovine GDF8 locus. *BMC Genetics*, 8, 80.

Komisarek, J., and Dorynek, Z. (2005). Polymorphisms of leptin and leptin receptor genes in the polish population of holstein-friesian bulls. *Annals of Animal Science*, 5, 253-260.

Konfortov, B.A., Licence, V.E and Miller, J.R. (1999). Re-sequencing of DNA from a diverse panel of cattle reveals a high level of polymorphism in both intron and exon. *Mammalian Genome*, 10, 1142-1145.

Konstantinos, K.V., Panagiotis, P., Antonios, V.T., Agelos, P., and Argiris, N.V. (2008). PCR-SSCP: A method for the molecular analysis of genetic diseases. *Molecular Biotechnology*, 38, 155-163.

Lagonigro, R., Wiener, P., Pilla, F., Woolliams, J.A., and Williams, J.L. (2003). A new mutation in the coding region of the bovine leptin gene associated with feed intake. *Animal Genetics*, 34, 371-374.

Laud, K., Gourdou, I., Belair, L., Keisler, D.H., and Djiane, J. (1999). Detection and regulation of leptin receptor mRNA in ovine mammary epithelial cells during pregnancy and lactation. *Federation of European Biochemical Societies Letters,* 463, 194-198.

Li, Y., Huitong, Z., Long, C., Grant, R.E., Jonathan, G.H.H. (2019). Effect of DGAT1 variant (K232A) on milk traits and milk fat composition in outdoor pasture-grazed dairy cattle. *New Zealand Journal of Agricultural Research,* DOI: 10.1080/00288233.2019.1589537.

Liefers, S. C., Te Pas, M.F.W., Veerkamp, R.F., and Van der Lende, T. (2002). Associations between leptin gene polymorphisms and production, live weight, energy balance, feed intake, and fertility in Holstein heifers. *Journal of Dairy Science*, 85, 1633-1638.

Liefers, S.C., Veerkamp, R.F., Te Pas, M.F.W., Delavaud, C., Chilliard, Y., Van der Lende, T. (2003). Fluctuations in leptin levels during late pregnancy and lactation in relation to dry matter intake, milk yield, live weight, energy balance and fertility in Holstein heifers. *Journal of Dairy Sci*ence, 86, 799-807.

Liefers, S.C., Veerkamp, R.F., te Pas, M.F.W., Delavaud, C., Chilliard, Y., Platje, M., and van der Lende, T. (2005). Leptin promoter mutations affect leptin levels and performance traits in dairy cows. *Animal Genetics*, 36, 111-118.

Lin, J., Barb, C.R., Matteri, R.L., Kraeling, R.R., Chen, X., Meinersmann, R.J., and Rampacek, G.B. (2000). Long form leptin receptor mRNA expression in the brain, pituitary, and other tissues in the pig. *Domestic Animal Endocrinolology*, 19, 53-61.

Lindersson, M., Andersson-Eklund, L., de Koning, D.J., Lunden, A., Maki-Tanila, A., and Andersson, L. (1998). Mapping of serum amylase-1 and quantitative trait loci for milk production traits to cattle chromosome 4. *Journal of Dairy Science*, 81, 1454-1461.

Liu, K., Sandgren, E.P., Palmiter, R.D., and Stein, A. (1995) Rat growth hormone gene introns stimulate nucleosome alignment in vitro and in transgenic mice. *Proceedings of the National Academy of Sciences,* USA 92, 7724-7728.

Loftus, R.T., Machugh, D.E., Bradley, D.G., Sharp, P.M., and Cunningham, E.P. (1994a). Evidence for two independent domestications of cattle. *Proceedings of the National Academy of Sciences,* USA 91, 2757-2761.

Lopez-Villalobos, N., and Garrick, D.J. (1996). Profitability of rotational crossbreeding programmes in commercial dairy herds. *Proceedings of the New Zealand Society of Animal Production,* 56.

Lowe, T. M., and Eddy, S.R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research*, 25, 955-964.

Madeja, Z., Adamowicz, T., Chmurzynska, A., Jankowski, T., Melonek, J., Switonski, M., Strabel, T. (2004). Short communication: effect of leptin gene polymorphisms on breeding value for milk production traits. *Journal of Dairy Science*, 87:3925-3927.

Marchitelli, C., Contarini, G., De Matteis, G., Crisa, A., Pariset, L., Scata, M.C., Catillo, G., Napolitano, F., and Moioli, B. (2013). Milk FA variability: Effect of some candidate genes involved in lipid synthesis. *Journal of Dairy Research*, 80, 165-173.

Maffei, M., Halaas, J. L., Ravussin, E., Pratley R.E., Lee, G. H., Zhang, Y., Fei, H., Kim, S., Lallone, R., Ranganathan, S., Kern, P.A., and Friedman, J.M. (1995). Leptin levels in human and rodent: Measurement of plasma leptin and ob RNA in obese and weight-reduced subjects. *Nature Medicine,* 1, 1155-1161.

Manickam, R., Pena, R.N., and Whitelaw, C.B.A. (2008) mammary gland differentiation inversely correlates with gdf-8 expression. *Molecular Reproduction and Development*, 75, 1783-1788.

Marie, M., Findlay, P.A., Thomas, L., and Adam, C.L. (2001). Daily patterns of plasma leptin in sheep: effects of photoperiod and food intake. *Journal of Endocrinology*, 170, 277-286.

Massague, J. (1998). TGF-b signal transduction. *Annual Review of Biochemistry,* 67, 753-791.

Mattick, J.S. (2001) Non-coding RNAs: the architects of eukaryotic complexity. *EMBO Reports,* 2, 986-991.

Mattioli, R.C., Pandey, V.S., Murray, M., and Fitzpatrick, J.L. (2000). Immunogenetic influences on tick resistance in African cattle with particular reference to trypanotolerant N'Dama (*Bos taurus*) and trypanosusceptible Gobra zebu (*Bos indicus*) cattle. *Acta Tropica*, 75, 263-277.

Maurice-Van Eijndhoven, M.H.T., Hiemstra, S.J., and Calus, M.P.L. (2011) Short communication: milk fat composition of 4 cattle breeds in the Netherlands. *Journal of Dairy Science*, 94, 1021-1025.

Meadow, R. H., (1993). Animal domestication in the Middle East: a revised view from the Eastern Margin, Harappan Civilization, Oxford & IBH, New Delhi, India, 295-320.

Mendias, C.L., Bakhurin, K.I., and Faulkner, J. A. (2008). Tendons of myostatin-deficient mice are small, brittle, and hypocellular. *Proceedings of the National Academy of Science*, USA, 105, 388-393.

Meuwissen, T.H.E., and Goddard, M.E. (1996). The use of marker haplotypes in animal breeding schemes. *Genetics Selection and Evolution*, 28, 161-176.

Muir, W.M. (2007). Comparison of genomic and traditional BLUP-estimated breeding value accuracy and selection response under alternative trait and genomic parameters. *Journal of Animal Breeding and Genetics*, 124, 342-355.

McMahon, C.D., Radcliff, R.P., Lookingland, K.J., and Tucker, H.A. (2001). Neuroregulation of growth hormone secretion in domestic animals. *Domestic Animal Endocrinology*, 20, 65-87.

McPherron, A.C, Lawler, A.M, and Lee, S.J. (1997). Regulation of skeletal muscle mass in mice by a new TGF-b superfamily member. *Nature,* 387, 83-90.

McPherron, A.C., and Lee, S.J. (1997). Double muscling in cattle due to mutations in the myostatin gene. *Proceedings of the National Academy of Sciences*, USA 94(23), 12457-12461.

McPherron, A.C., and Lee, S.J. (2002). Suppression of body fat accumulation in myostatin-deficient mice. *The Journal of Clinical Investigation,* 109, 595-601.

Mele, M., Macciotta, N.P.P., Cecchinato, A., Conte, G., Schiavon, S., and Bittante, G. (2016). Multivariate factor analysis of detailed milk FA profile: Effects of dairy system, feeding, herd, parity, and stage of lactation. *Journal of Dairy Science*, 99(12), 9820-9833.

Moran, J. (2015) Managing cow lactation cycles, the cattle site. http://www.thecattlesite.com/articles/4248/managing-cow-lactation-cycles/ (accessed 27th October, 2019).

Morrison, C.D., Daniel, J.A., Holmberg, B. J., Djiane, J., Raver, N., Gertler, A., and Keisler, D.H. (2001). Central infusion of leptin into well-fed and undernourished ewe lambs: effects on feed intake and serum concentrations of growth hormone and luteinizing hormone. *Journal of Endocrinology*, 168, 317-324.

Mosher, D.S., Quignon, P., Bustamante, C.D., Sutter, N.B., Mellersh, C.S., Parker, H.G., and Ostrander, E.A. (2007). A mutation in the myostatin gene increase muscle mass and enhances racing performance in heterozygote dogs. *PLoS Genetics*, 3, 779-786.

Mounzih, K., Qiu, J., Ewart-Toland, A., and Chehab, F.F. (1998). Leptin is not necessary for gestation and parturition but regulates maternal nutrition via a leptin resistance state. *Endocrinology*, 139, 5259-5262.

Moussavi, A.H., Ahouei, M., Nassiry, M.R., and Javadmanesh, A. (2006). Association of leptin polymorphism with production, reproduction and plasma glucose level in Iranian Holstein Cows. Asian-Austral. *Journal of Animal Science*, 19, 627-631.

Mukherjea, R., T. W. Castonguay, L. W. Douglass, and P. Moser-Veillon. (1999). Elevated leptin concentrations in pregnancy and lactation: possible role as a modulator of substrate utilization. *Life Science*, 65, 1183-1193.

Nafikov, R.A., Schoonmaker, J.P., Korn, K.T., Noack, K., Garrick, D.J., Koehler, K. J., Minick-Bormann, J., Reecy, J.M., Spurlock, D.E., Beitz, D. C. (2013). Association of polymorphisms in solute carrier family 27, isoform A6 (SLC27A6) and FA-binding protein-3 and FA-binding protein-4 (FABP3 and FABP4) with FA composition of bovine milk. *Journal of Dairy Science*, 96(9), 6007-6021.

Ohshiro, Y., Ueda, K., Nishi, M., Ishigame, M., Wakasaki, H., Kawashima, H., Furuta, H., Sasaki, H., Sanke, T., Takasu, N., and Nanjo, K. (2000). A polymorphic marker in the leptin gene associated with Japanese morbid obesity. *Journal of Molecular Medicine*, 78, 516-520.

Orbak, Z., Coker, M., Darcan, S., and Goksen, D. (2001). Association between serum leptin and anthropometric parameters at birth and at 15th day of life in infants born asymmetrically small for gestational age. *Journal of Pediatric Endocrinology and Metabolism*, 14, 185-192.

Ovilo, C., Clop, A., Noguera, J.L., Oliver, M. A., Barragan, C., Rodriguez, C., Silio, L., Toro, M. A., Coll, A., Folch, J. M., Sanchez, A., Babot, V., Varona, L., and Perez-Enciso, M. (2002). Quantitative trait locus mapping for meat quality traits in an Iberian x Landrace F2 pig population. *Journal of Animal Science*, 80, 2801-2808.

Palladino, R.A., Buckley, F., Prendiville, R., Murphy, J.J., Callan, J., and Kenny, D.A. (2010). A comparison between Holstein-Friesian and Jersey dairy cows and their f1 hybrid on milk FA composition under grazing conditions. *Journal of Dairy Science,* 93, 2176-2184.

Palmquist, D.L., Beaulieu, A.D., Barbano, D.M. (1993). Feed and animal factors influencing milk fat composition. *Journal of Dairy Science*, 76, 1753-1771.

Pegolo, S., Cecchinato, A., Casellas, J., Conte, G., Mele, M., Schiavon, S., and Bittante, G. (2016). Genetic and environmental relationships of detailed milk FAs profile determined by gas chromatography in Brown Swiss cows. *Journal of Dairy Science*, 99, 1315-1330.

Pegolo, S., Cecchinato, A., Mele, M., Conte, G., Schiavon, S., and Bittante, G. (2016). Effect of candidate gene polymorphisms on the detailed FAs profile determined by gas chromatography in bovine milk. *Journal of Dairy Science*, 99, 4558-4573.

Penno, J. W., Macdonald, K. A., Holmes, C. W., Davis, S. R., Wilson, G. F., Brookes, I. M., and Thom, E. R. (2007). Responses to supplementation by dairy cows given low pasture allowances in different seasons 1. Pasture intake and substitution. *Animal Science*, 82(5), 661-670.

Pickavance, L., Tadayyon, M., Williams, G., and Vernon, R.G. (1998). Lactation suppresses diurnal rhythm of serum leptin. *Biochemical and Biophysical Resource Commun*ication, 248, 196-199.

Pirlo, G., Miglior, F., Speroni, M. (2000). Effect of age at first calving on production traits and on difference between milk yield returns and rearing costs in Italian Holsteins. *Journal Dairy Science*, 83:603-608.

Pomp, D., Zou, T., Clutter, A. C., and Barendse, W. (1997). Rapid communication: mapping of leptin to bovine chromosome 4 by linkage analysis of a PCR-based polymorphism. *Journal of Animal Sci*ence, 75, 1427.

Raes, K., de Smet, S., and Demeyer, D. (2001). Effect of double-muscling in Belgian Blue young bulls on the intramuscular FA composition with emphasis on conjugated linoleic acid and polyunsaturated FAs. *Animal Science,* 73, 253-260.

Ren, M.Q., Wegner, J., Bellmann, O., Brockmann, G.A., Schneider, F., Teuscher, F., and Ender, K. (2002). Comparing mRNA levels of genes encoding leptin, leptin receptor, and lipoprotein lipase between dairy and beef cattle. *Domestic Animal Endocrinology*, 23, 371-381.

Richert, M.M., Schwertfeger, K.L., Ryder, J.W., and Anderson, S.M. (2000) .An atlas of mouse mammary gland development. *Journal of Mammary Gland Biology and Neoplasia,* 5, 227-241.

Robinson, S.D., Silberstein, G.B., Roberts, A.B., Flanders, K.C., and Daniel, C.W. (1991). Regulated expression and growth inhibitory effects of transforming growth factor bisoforms in mouse mammary gland development. *Development*, 113, 867-878.

Robinson, G.W., McKnight, R.A., Smith, G.H., and Hennighausen, L. (1995). Mammary epithelial cells undergo secretory differentiation in cycling virgins but require pregnancy for the establishment of terminal differentiation. *Development,* 121, 2079-2090.

Robinson, G.W., Karpf, A.B., Kratochwil, K. (1999). Regulation of mammary gland development by tissue interaction. *Journal of Mammary Gland Biology and Neoplasia,* 4, 9-19.

Roche, J.R., Berry, D.P., and Kolver, E.S. (2006). Holstein-Friesian strain and feed effects on milk production, body weight, and body condition score profiles in grazing dairy cows. *Journal of Dairy Science,* 89(9), 3532-3543.

Rock, F.L., Altmann, S.W., Van Heek, M., Kastelein, R.A., and Bazan, J.F. (1996). The leptin haemopoietic cytokine fold is stabilized by an intrachain disulfide bond. *Hormorne and Metabolic Resource*, 28, 649-652.

Santos-Alvarez, J., Goberna, R., and Sa´nchez-Margalet, V. (1999). Human leptin stimulates proliferation and activation of human circulating monocytes. *Cellular Immunology*, 194, 6-11.

Sari, A. (2013). The relationship between leptin and FA. *Journal of Molecular Biomarkers and Diagn*osis, 4, 139.

Sauna, Z.E., Kimchi-Sarfaty, C. (2011). Understanding the contribution of synonymous mutations to human disease. *Nature Reviews Genetics*, 12, 683-691.

Schrooten, C., Bovenhuis, H., Coppieters, W., and Van Arendonk, J. A. (2000). Whole genome scan to detect quantitative trait loci for conformation and functional traits in dairy cattle. *Journal of Dairy Science,* 83, 795-806.

Schuelke, M., Wagner, K.R., Stolz, L.E., Hübner, C., Riebel, T., Kömen, W., Braun, T., Tobin, J.F., and Lee, S.J. (2004). Myostatin mutation associated with gross muscle hypertrophy in a child. *The New England Journal of Medicine*, 350, 2682-2688.

Seeber, R.M., Smith, J.T., and Waddell, B.J. (2002). Plasma leptin-binding activity and hypothalamic leptin receptor expression during pregnancy and lactation in the rat. *Biology of Reproduction*, 66, 1762-1767.

Serjsen, K. (2005). Mammary development. In Calf and Heifer Rearing: Principles of Rearing the Modern Dairy Heifer from Calf to Calving. First edn. Ed P. C. Garnsworthy. Nottingham University Press 237-251.

Sjakste, T., Paramonova, N., Grislis, Z., Trapina, I., and Kairisa, D. (2011). Analysis of the single-nucleotide polymorphism in the 5'UTR and part of intron 1 of the sheep MSTN gene. *DNA and Cell Biology*, 30, 433-444.

Smith, T. P., Lopez-Corrales, N. L., Kappes, S. M., and Sonstegard, T. S. (1997). Myostatin maps to the interval containing the bovine mh locus. *Mammalian Genome*, 8(10), 742-744.

Smith, J.A., Lewis, A.M., Wiener, P., and Williams, J.L. (2000). Genetic variation in the bovine myostatin gene in UK beef cattle: allele frequencies and haplotype analysis in the South Devon. *Animal Genetics*, 31, 306-309.

Smith, J. L., and Sheffield, L.G. (2002). Production and regulation of leptin in bovine mammary epithelial cells. *Domestic Animal Endocrinology*, 22, 145-154.

Sonstegard, T.S., Rohrer, G.A., and Smith, T.P.L. (1998). Myostatin maps to Chromosome 15 by linkage and physical analyses. *Animal Genetics*, 29, 19-22.

Soyeurt, H., Dardenne, P.F., Dehareng, G., Lognay, G., Veselko, D., Marlier, M., Bertozzi, C., Mayeres, P., and Gengler, N. (2006). Estimating FA content in cow milk using mid-infrared spectrometry. *Journal of Dairy Science,* 89, 3690-3695.

Sprang S.R., Bazan J.F., 1993. Cytokine structural taxonomy and mechanisms of receptor engagement. *Current Opinion in Structural Biology*, 5, 114-121.

Stelwagen, K. (2011). Mammary Gland, Milk Biosynthesis and Secretion | Lactose. In J. W. Fuquay (Ed.), Encyclopedia of Dairy Sciences, San Diego: Academic Press, 367-372.

Stone, R.T., Kappes, S.M., and Beattie, C.W. (1996b). Two polymorphic microsatellites within an 18 kb genomic clone containing the bovine ob gene. *Animal Genetics*, 27, 64.

Stoop, W.M., Van Arendonk, J.A.M., Heck, J.M.L., Van Valenberg, H.J.F., Bovenhuis, H. (2008). Genetic parameters for major milk FAs and milk production traits of Dutch Holstein Friesians. *Journal of Dairy Science*, 91, 385-394.

Stoop, W.M., Bovenhuis, H., Heck, J.M.L., and Van Arendonk, J.A.M. (2009). Effect of lactation stage and energy status on milk fat composition of Holstein-Friesian cows. *Journal of Dairy Science*, 92(4), 1469-1478.

Strucken, E.M., Laurenson, Y.C.S.M., and Brockmann, G.A. (2015). Go with the flow-biology and genetics of the lactation cycle. *Frontiers in Genetics*, 6, 118.

Sutton, J.D. (1989). Altering milk composition by feeding. Journal of Dairy Science, 72:2801-2814.

Tang-Christensen, M., Havel, P. J., Jacobs, R. R., Larsen, P. J. and Cameron, J.L. (1999). Central administration of leptin inhibits food intake and activates the sympathetic nervous system in rhesus macaques. *Journal of Clinical Endocrinology and Metabolism*, 84, 711-717.

Te Ara-the Encyclopaedia of NZ (accessed 16 November, 2018) https://teara.govt.nz/en

Tantia, M.S., Vijh, R.K., Kumar, S.T., Mishra, B., Reecy, J.M. (2006). Comparative analysis of GDF 8 (myostatin) in Bos indicus and Bos taurus. *DNA Seq*uence, 17, 311-313.

Thomas, M., Langley, B., Berry, C., Sharma, M., Kirk, S., Bass, J., and Kambadur, R. (2000). Myostatin, a negative regulator of muscle growth, functions by inhibiting myoblast proliferation. *Journal of Biological Chemistry*, 275, 40235-40243.

Thomas, L., Wallace, J. M., Aitken, R. P., Mercer, J. G., Trayhurn, P. and N. Hoggard. (2001). Circulating leptin during ovine pregnancy in relation to maternal nutrition, body composition and pregnancy outcome. *Journal of Endocrinology*, 169, 465-476

Tomimatsu, T., Yamaguchi, M., Murakami, T., Ogura, K., Sakata, M., Mitsuda, N., Kanzaki, T., Kurachi, H., Irahara, M., Miyake, A., Shima, K., Aono, T., and Murata, Y. (1997). Increase of mouse leptin production by adipose tissue after midpregnancy: gestational profile of serum leptin concentration. *Biochemical and Biophysical Research Communication*, 240, 213-215.

Trakovicka, A., Moravcikova, N., Minarovic, T., Alica, Navratilova, A. (2015). SNPs analyses of the bovine LEP and PIT-1 genes by multiplex PCR-RFLP method and their e_ect on milk performance traits in Slovak Simmental cattle. *Journal of Central European Agriculture*, 16, 65-75.

Vrankovic, L., Aladrovic, J., Octenjak, D., Bijelic, D., Cvetnic, L., and Stojevic, Z. (2017). Milk FA composition as an indicator of energy status in Holstein dairy cows. *Archives Animal Breeding*, 60, 205-212.

Wang, Z., and Burge, C.B. (2008). Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *RNA,* 14, 802-813.

Watanobe, H. (2002). Leptin directly acts within the hypothalamus to stimulate gonadotropin releasing hormone secretion in vivo in rats. *The Journal of Physiology*, 545, 255-268.

Whittemore, L.A., Song, K., Li, X., Aghajanian, J., Davies, M., Girgenrath, S., Hill, J.J., Jalenak, M., Kelley, P., Knight, A., Maylor, R., O'Hara, D., Pearson, A., Quazi, A., Ryerson, S., Tan, X.Y., Tomkinson, K.N., Veldman, G.M., Widom, A., Wright, J.F., Wudyka, S., Zhao, L., and Wolfman, N.M. (2003). Inhibition of myostatin in adult mice increases skeletal muscle mass and strength. *Biochemical and Biophysical Research Communications*, 300, 965-971.

Wiener, P., Woolliams, J.A., Frank-Lawale, A., Ryan, M., Richardson, R.I., Nute, G.R., Wood, J.D., Homer, D., and Williams, J.L. (2009). The effects of a mutation in the myostatin gene on meat and carcass quality. *Meat Science*, 83, 127-134.

Woller, M., Tessmer, S., Neff, D., Nguema, A. A., Roo, B.V. and Waechter-Brulla, D. (2001). Leptin stimulates gonadotropin releasing hormone release from cultured intact hemihypothalami and enzymatically dispersed neurons. *Experimental Biology and Medicone*, 226, 591-596.

Woodside, B., Abizaid, A. and Walker, C. (2000). Changes in leptin levels during lactation: implications for lactational hyperphagia and anovulation. *Hormornes and Behaviour*, 37, 353-365.

Winter, A., Krämer, W., Werner, F.A.O., Kollers, S., Kata, S., Durstewitz, G., Buitkamp, J.,Womack, J.E.,Thaller, G., and Fries, R. (2002). Association of a lysine-232/alanine polymorphism in a bovine gene encoding acyl-CoA: diacylglycerol acyltransferase (DGAT1) with variation at a quantitative trait locus for milk fat content. *Proceedings of the National Academy of Sciences*, 99(14), 9300-9305.

Yang, D., Chen, H., Wang, X., Tian, Z., Tang, L., Zhang, Z., Lei, C., Zhang, L., and Wang, Y. (2007). Association of Polymorphisms of *Leptin* Gene with Body Weight and Body Sizes Indexes in Chinese Indigenous Cattle. *Journal of Genetics and Genomics*, 34(5), 400-405

Yang, Y., Tang, B., Robinson, G., Hennighausen, L., Brodie, S.G., Deng, C-X., and Wakefield, L.M. (2002). Smad3 in the mammary epithelium has a nonredundant role in the induction of apoptosis, but not in the regulation of proliferation or differentiation by transforming growth factor-b. *Cell Growth and Differentiation*, 13, 123-130.

Ying, D., Sham, P.C., Smith, D.K., Zhang, L., Lau, Y.L., and Yang, W (2015). HaploShare: identification of extended haplotypes shared by cases and evaluation against controls. *Genome Biology* (Online Edition), 16, article no. 92 http://hdl.handle.net/10722/211876

Li, Y., Zhou, H., Cheng, L., Edwards, G.R., and Hickford, J.G.H. (2019). Effect of DGAT1 variant (K232A) on milk traits and milk fat composition in outdoor pasture-grazed dairy cattle. *New Zealand Journal of Agricultural Research.* DOI: 10.1080/00288233.2019.1589537

Yutzey, K.E., Kline, R.L., and Konieczny, S.F. (1989). An internal regulatory element controls troponin I gene expression. *Molecular and Cellular Biology*, 9, 1397-1405.

Zieba, D.A., Amstalden, M., Morton, S., Gallino, J. L., Edwards, J. F., Harms, P. G., and G. L. Williams. (2003a). Effects of leptin on basal and GHRH-stimulated GH secretion from the bovine adenohypophysis are dependent upon nutritional status. *Journal of Endocrinology*, 178, 83-89.

Zhang, Y., Proenca, R., Maffei, M., Barone, M., Leopold, L., and Friedman, J.M. (1994). Positional cloning of the mouse obese gene and its human homologue. *Nature*, 372, 425-432

Zhang, F., Basinski, M. B., Beals, J. M., Briggs, S. L., Churgay, L. M., Clawson, D. K., DiMarchi, R. D., Furman, T. C., Hale, J. E.,. Hsiung, H. M., Schoner, B. E., Smith, D. P., Zhang, X. Y., Wery, J. P., and Schevitz, R. W. (1997). Crystal structure of the obese protein leptin-E100. *Nature*, 387, 206-209.

Zhou, H., Hickford, J.G.H, and Fang, Q. (2006). A two-step procedure for extracting genomicDNA from dried blood spots on filter paper for polymerase chain reaction amplification. *Analytical Biochemistry*, 354, 159-161

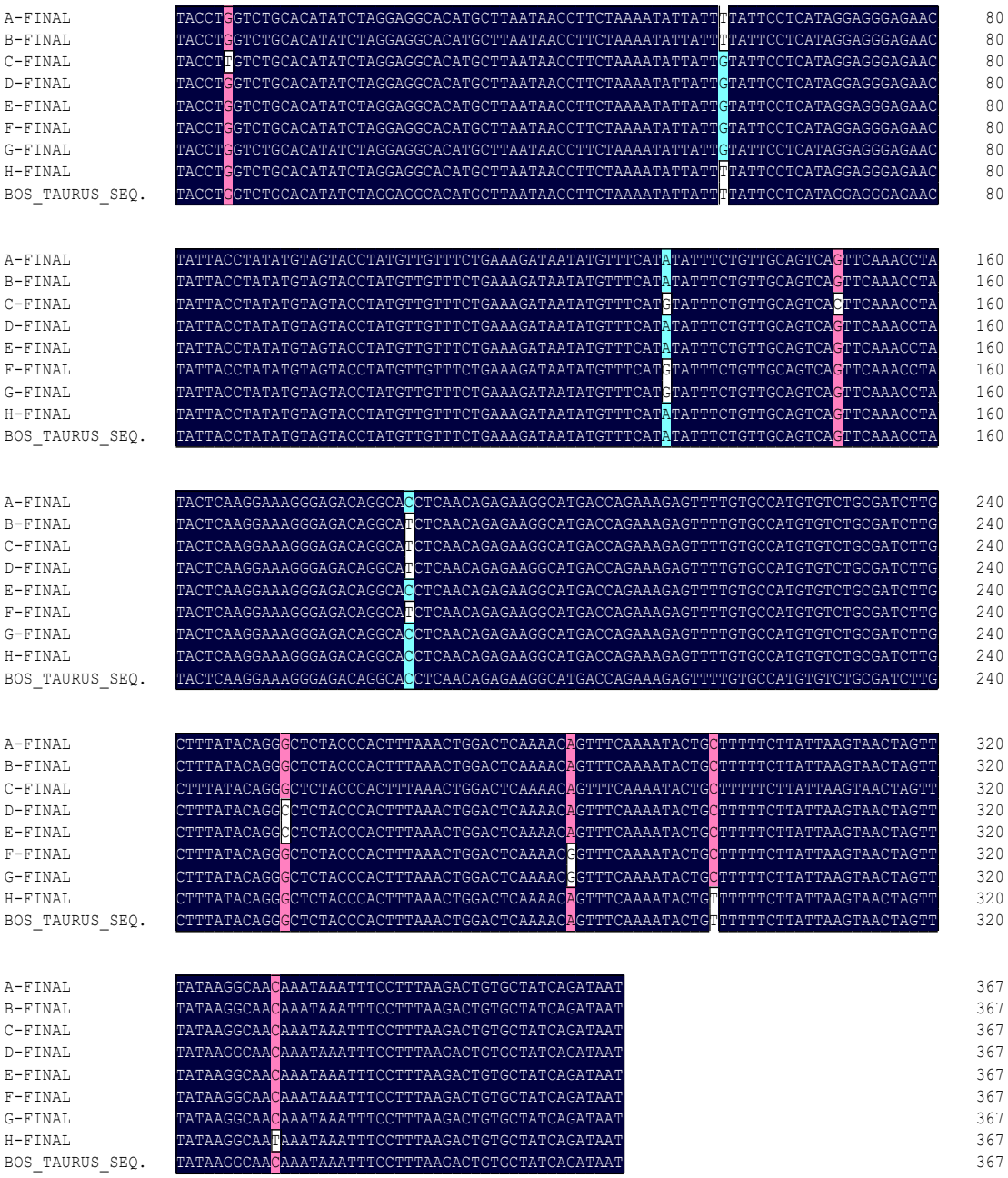
Zhou, H., Hickford, J.G.H., and Fang, Q. (2008). Variation in the coding region of the myostatin (GDF8) gene in sheep. *Molecular and Cellular Probes*, 22, 67-68.

Zhu, X., Topouzis, S., Liang, L-F., and Stotish, R.L. (2004). Myostatin signaling through Smad2, Smad3 and Smad4 is regulated by the inhibitory Smad7 by a negative feedback mechanism. *Cytokine.* 26, 262-272.

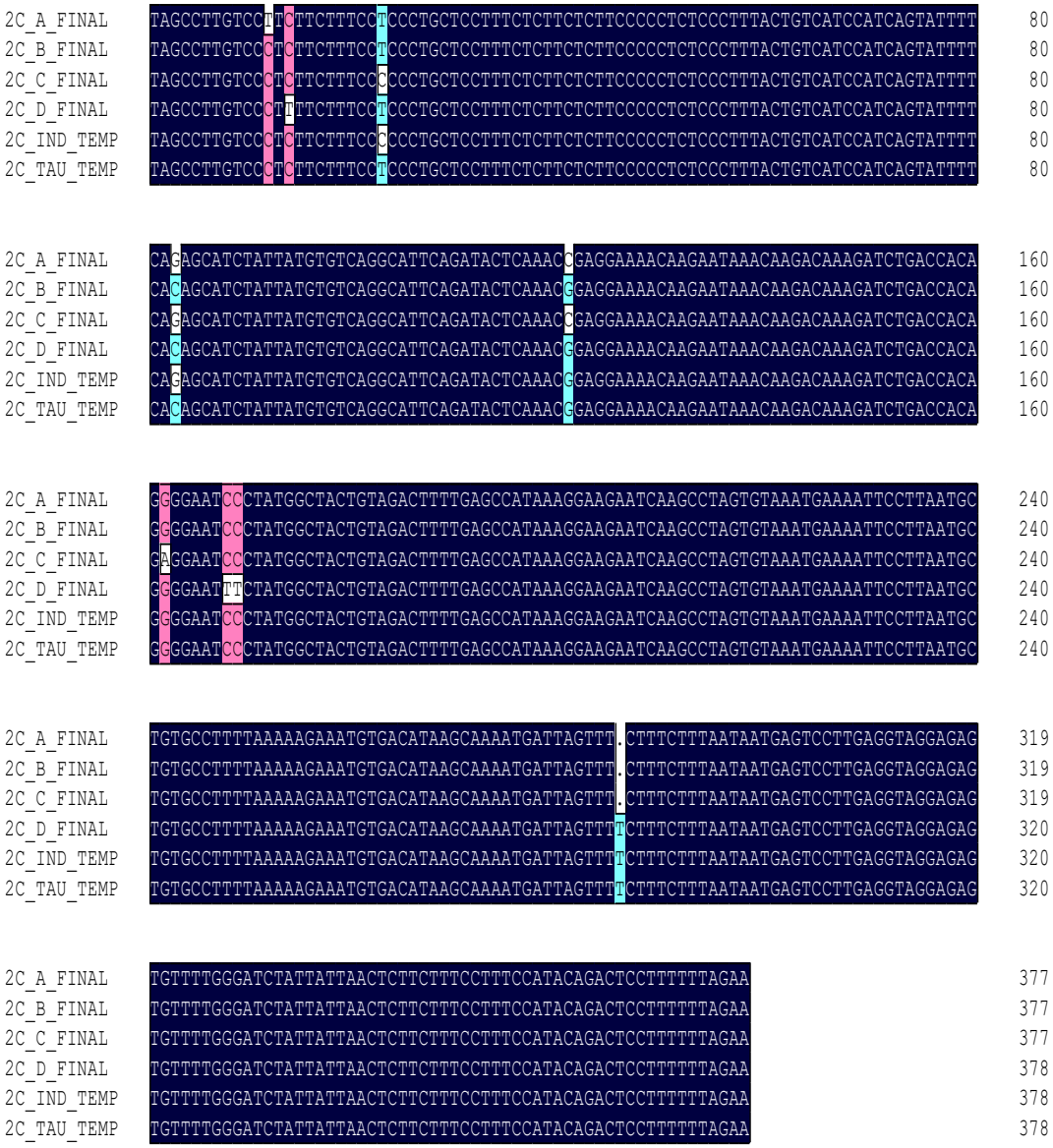
Appendix A

Bovine Myostatin gene sequence

A.1 *MSTN* intron 1 region examined



A.2 *MSTN* intron 2 region examined



Appendix B

Bovine leptin gene sequence

B.1 *LEP* intron 1/exon 2 region

A_1 Variant sequence

GTCTTTGAGGAGATGATAGCCATGGCAGACAGCAAATCTCGTTGTTATCCGCATCTGAAG
ACGTGGATGCGGGTGGTAACGGAGCACGTGGGTGTTCTCGGAGATCGACGATGTGCCAC
GTGTGGTTTCTTCTGTTTTCAGGCCCCAGAAGCCCATCCCGGGAAGGAAAATGCGCTGTG
GACCCCTGTATCGATTCCTGTGGCTTTGGCCCTATCTGTCTTACGTGGAGGCTGTGCCCAT
CCGCAAGGTCCAGGATGACACCAAAACCCTCATCAAGACAATTGTCACCAGGATCAATG
ACATCTCACACACGGTAGGGAGGGACTGGGAGACGAGGTAGAACCGTGGCCATCCCGTG
GGGGACCCCAGAGGCTGGCGGAGGAGGCTGTGCAGCCTTGCACAGGGCCCCAGTGGCCT
GGACGCCCCCCTGGC

B_1 Variant sequence

CAACAAGATTTGCTTGTCTTTGAAGGGATGATAGCCATGGCAGACAGCAAATCTTGTTGT
TATCCGCATCTGAAGACCTGGATGCGGGTGGTAACGGAGCACGTGGGTGTTCTCGGAGA
TCGACGATGTGCCACGTGTGGTTTCTTCTGTTTTCAGGCCCCAGAAGCCCATCCCGGGAA
GGAAAATGCGCTGTGGACCCCTGTATCGATTCCTGTGGCTTTGGCCCTATCTGTCTTACGT
GGAGGCTGTGCCCATCTGCAAGGTCCAGGATGACACCAAAACCCTCATCAAGACAATTG
TCACCAGGATCAATGACATCTCACACACGGTAGGGAGGGACTGGGAGACGAGGTAGAAC
CGTGGCCATCCCGTGGGGGACCCCAGAGGCTGGCGGAGGAGGCTGTGCAGCCTTGCACA
GGGCCCCAGTGGCCTGGACGCCCCCCTGGAA

C_1 Variant sequence

ACAACAAGATTTGCTTGTCTTTGAGGAGATGATAGCCATGGCAGACAGCAAATCTTGTTG
TTATCCGCATCTGAAGACCTGGATGCGGGTGGTAACGGAGCACGTGGGTGTTCTCGGAG
ATCGACGATGTGCCACGTGTGGTTTCTTCTGTTTTCAGGCCCCAGAAGCCCATCCCGGGA
AGGAAAATGCGCTGTGGACCCCTGTATCGATTCCTGTGGCTTTGGCCCTATCTGTCTTAC
GTGGAGGCTGTGCCCATCCGCAAGGTCCAGGATGACACCAAAACCCTCATCAAGACAAT
TGTCACCAGGATCAATGACATCTCACACACGGTAGGGAGGGACTGGGAGACGAGGTAGA
ACCGTGGCCATCCCGTGGGGGACCCCAGAGGCTGGCGGAGGAGGCTGTGCAGCCTTGCA
CAGGGCCCCAGTGGCCTGGACGCCCCCCTGGCA

D_1 Variant sequence

GATGATAGCCATGGCAGACAGCAAATCTCGTTGTTATCCGCATCTGAAGACGTGGATGC
GGGTGGTAACGGAGCACGTGGGTGTTCTCGGAGATCGACGATGTGCCACGTGTGGTTTCT
TCTGTTTTCAGGCCCCAGAAGCCCATCCCGGGAAGGAAAATGCGCTGTGGACCCCTGTAT
CGATTCCTGTGGCTTTGGCCCTATCTGTCTTACGTGGAGGCTGTGCCCATCCGCAAGGTCC
AGGATGACACCAAAACCCTCATCAAGACAATTGTCACCAGGATCAATGACATCTCACAC
ACGGTAGGGAGGGACTGGGAGACGAGGTAGAACCGTGGCCATCCGTGGGGGACCCCAG
AGGCTGGCGGAGGAGGCTGTGCAGCCTTGCACAGGGCCCCAGTGGCCTGGACGCCCCCC
TGGT

B.2 *LEP* intron 2 region

A_2 Variant sequence

CAAGGTTCCACAAGGTTTTTAAACTCCAGTTTCCTCATCTAGAAAATGAAAGTGGGAAAG
TGTTAGTTGCTCAGTCATGTCCAACTCTTTGAGACCCCATGAACTGTAGTCTACCAGGCTC
CTCTGTCCATGAAATTCTTCAGGCAAGAATACTGGAGTGGCTTGTTATTTTCTTCTCCCAA
CAAGATCTTCCCAACCCAGGGATTGAACCTGGGTCTTCTAAATTGCAGGCAGATTCTTTA
CCGTCTGAGCCACCAGGGAAACCCATAAGACCTTGTGAAGACTATTAAGATAGTCATCT
AGACAACAAGACTATCTTAATAGTCTTCATAAGGTCTTCATGAGACTAAATTAGATAAAG
CAAGTGACCCTCCTGAATACCTTTTGCAGAACCCAGGTTCAATCTATGGGTTGGGAAGAT
CTTGTTGGGAGAAGAAAATAACAAGCCACTCCAGTATTCTTGCCTGAAGAATTTCATGGA
CAGAGGAGC

B_2 Variant sequence

TCAGGTTCCACAAGGTTTTTAAACTCCAGTTTCCTCATCTAGAAAATGAAAGTGGGAAAG
TGTTAGTTGCTCAGTCATGTCCAACTCTTTGAGACCCCATGAACTGTAGTCTACCAGGCTC
CTCTGTCCATGAAATTCTTCAGGCAAGAATACTGGAGTGGCTTGTTATTTTCTTCTCCCAA
CAAGATCTTCCCAACCCAGGGATTGAACCTGGGTCTTCTAAATTGCAGGCAGATTCTTTA
CCGTCTGAGCCACCAGGGAAACCCATAAGATCTTGTGAAGACTATTAAGATAGTCATCTA
GACAACAAGACTATCTTAATAGTCTTCATAAGGTCTTCATGAGACTAAATTAGATAAAGC
AAGTGACCCTCCTGAATACC

C_2 Variant sequence

AGCTAGTCCAGTTCCACAAGGTTTTTAAACTCCAGTTTCCTCATCTAGAAAATGAAAGTG
GGAAAGTGTTAGTTGCTCAGTCATGTCCAACTCTTTGAGACCCCATGAACTGTAGTCTAC
CAGGCTCCTCTGTCCATGAAATTCTTCAGGCAAGAATACTGGAGTGGCTTGTTATTTTCTT
CTCCCAACAAGATCTTCCCAACCCAGGGATTGAACCTGGGTCTTCTAAATTGCAGGCAGA
TTCTTTACCGTCTGAGCCACCAGGGAAACCCATAAGACCTTGTGAAGACTATTAAGATAG
TCATCTAGACAACAAGACTATCTTAATAGTCTTCATAAGGTCTTCATGAGACTAAATTAG
ATAAAGCAAGTGACCTCCCTGAATACC

B.3 *LEP* intron 2/exon 3 region

A_3 Variant sequence

TTCGGAAGGCAGACTGGGAGGATCGTTGTAGATCGCCAAGTCTGTCCATCTGGACAACTC
AGGAGAGGGGAGCCAGGTGAAGTCCAAACCAGTGACCCCGTTTGCTCTCCCCTTCCTCCT
GCATAGCAGTCCGTCTCCTCCAAACAGAGGGTCACTGGTTTGGACTTCATCCCTGGGCTC
CACCCTCTCCTGAGTTTGTCCAAGATGGACCAGACATTGGCGATCTACCAACAGATCCTC
ACCAGTCTGCCTTCCAGAAATGTGGTCCAAATATCCAATGACCTGGAGAACCTCCGGGAC
CTTCTCCACCTGCTGGCCGCCTCCAAGAGCTGCCCCTTGCCGCAGGTCAGGGCCCTGGAG
AGCTTGGAGAGCTTGGGCGTTGTCCTGGAAGCTTCCCTCTACTCCACCGAGGTGGTGGCC
CTGAGCCGGCTGCAGGGGTCACTACAGGACATGTTGCGGCAGCTGGACCTCAGTCCCGG
GTGCTGAAGCCTTGAAGGCCTCTCTTCCCAAAGTCCAGGGAAA

B_3 Variant sequence

CGGGATGAAGTCCAAACCAGTGACCCTACGTTTTGCTCTCCCCTTCCTCCTGCATAGCAG
TCCGTCTCCTCCAAACAGAGGGTCACTGGTTTGGACTTCATCCCTGGGCTCCACCCTCTCC
TGAGTTTGTCCAAGATGGACCAGACATTGGTGATCTACCAACAGATCCTCACCAGTCTGC
CTTCCAGAAATGTGGTCCAAATATCCAATGACCTGGAGAACCTCCGGGACCTTCTCCACC
TGCTGGCCGCCTCCAAGAGCTGCCCCTTGCCGCAGGTCAGGGCCCTGGAGAGCTTGGAG
AGCTTGGGCGTTGTCCTGGAAGCTTCCCTCTACTCCACCGAGGTGGTGGCCCTGAGCCGG
CTGCAGGGGTCACTACAGGACATGTTGCGGCAGCTGGACCTCAGTCCCGGGTGCTGAAG
CCTTGAAGGCCTCTCTTCCCAAAGTCCAGGGAAG

C_3 Variant sequence

GTTTGCTCTCCCCTTCCTCCTGCATAGCAGTCCGTCTCCTCCAAACAGAGGGTCACTGGTT
TGGACTTCATCCCTGGGCTCCACCCTCTCCTGAGTTTGTCCAAGATGGACCAGACATTGG
CGATCTACCAACAGATCCTCACCAGTCTGCCTTCCAGAAATGTGGTCCAAATATCCAATG
ACCTGGAGAACCTCCGGGACCTTCTCCACCTGCTGGCCGCCTCCAAGAGCTGCCCCTTGC
CGCAGGTCAGGGCCCTGGAGAGCTTGGAGAGCTTGGGTGTCGTCCTGGAAGCCTCCCTCT
ACTCCACCGAGGTGGTGGCCCTGAGCCGGCTGCAGGGGTCACTACAGGACATGTTGCGG
CAGCTGGACCTCAGTCCTGGGTGCTGAAGCCTTGAAGGCCTCTCTTCCCAAAGTCCAGGG
AAA